THE HONG KONG
POLYTECHNIC UNIVERSITY
香港理工大學
Pao Yue-kong Library
包玉剛圖書館

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

Pao Yue-kong Library, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

http://www.lib.polyu.edu.hk

THE HONG KONG POLYTECHNIC UNIVERSITY

DEPARTMENT OF LAND SURVEYING AND GEO-INFORMATICS

AUTOMATIC FUSION OF PHOTOGRAMMETRIC IMAGERY AND LASER SCANNER POINT CLOUDS

ERIC KWABENA FORKUO

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

March, 2005

Pao Yue-kong Library
PolyU · Hong Kong

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

(Signed)

Eric Kwabena Forkuo (Name of Student)

This thesis is dedicated to my wife Gloria and my sons Jesse and Obed. Without their love, support, encouragement and understanding this thesis would never have been completed.

ABSTRACT

Close-range photogrammetry and the relatively new technology of terrestrial laser scanning can be considered as complementary rather than competitive technologies. For instance, terrestrial laser scanners (TLS) have the ability to rapidly collect high-resolution 3D surface information about an object. The same type of data can be generated using close-range photogrammetric (CRP) techniques, but image disparities common to close-range scenes makes this an operator intensive task. The imaging systems of some TLSs do not have very high radiometric resolution whereas high-resolution digital cameras used in modern CRP do. Finally, TLSs are essentially earth-bound whereas cameras can be moved at will around the object being imaged.

This thesis, therefore, explores and attempts to provide a solution to the problems of developing a methodology to fuse terrestrial laser scanner generated 3D data and high-resolution digital images. Four phases of the methodology have been investigated:- data pre-processing (fusion of data from the two sensors), automatic measurements (feature detection and correspondence matching), mapping (creation of point cloud visual index), and orientation (calculation of exterior orientation parameters). Individual phases were initially investigated in a manually controlled environment, typically using commercial photogrammetric software, and then combined in a completely automated system.

Focusing on the amount of geometric primitives, three different scenes (data set A, data set B, and data set C) representing three levels of complexity (low, medium and high) were scanned with the laser scanner, and for each scan, a 2D photographic image was taken with a digital camera. To overcome the differences in datasets, a hybrid matching (both feature and area-based) algorithm was successfully developed and implemented.

The fidelity of the concept of generating synthetic camera images has been tested by determining the exterior orientation of the synthetic camera images and the real camera images relative to the point cloud. This orientation process was first achieved by using manual methods and existing photogrammetric application software.

The results verified that there were no conceptual errors in the developed methods. However, in order to meet the objective of this thesis, an automatic technique with photogrammetric bundle adjustment was developed. Three different sets of data were used to check the validity and reliability of the developed methodology. The results of measurements on interest points and correspondence matching are presented. Also, the results of manual and automatic exterior orientation are presented. The results indicate that the concept of the synthetic camera image is a feasible method for multisensor fusion. The greatest promise is offered by the point cloud visual index.

ACKNOWLEDGMENTS

I would like to express my sincere appreciation and indebtedness to my thesis supervisor Dr. Bruce King for his valuable suggestion, guidance and supervision towards the successful accomplishment of this research work. I would like to express my thanks to my co-supervisor Professor Gold for his guidance, encouragement and constant support over the years. Without his encouragement, studying in this department would not be possible. I would also like to give a special thanks to his wife, Valerie for her valuable moral support, love and kindness. To them I am particularly grateful.

Special thanks to Prof. Michael Chapman (Ryerson University), Dr. Bruce Harvey (University of New South Wales) and Prof. X. L. Ding for serving on my examining committee and for their helpful recommendations. My thanks are extended to Prof. Zhilin Li for his patience in discussing my questions and the comments on the contents of the thesis. I also express my sincere gratitude to Nelson Chan for his help and support through my doctoral study. Then I would like to thank Dr. Gillian Humphreys for proofreading my thesis and offering critical evaluations. I would like to express my gratitude to Dr. Janet Nichol for her support.

Dr. Vijay Vohora, Peter Damoah-Afari, Charles Marfo and Asma Tradu Ibrahim not only encouraged me but supported me and offered very helpful critical

evaluations. To them I am particularly grateful their warm friendship and assistance in both professional and personal perspectives.

I owe a lot to my friends Pieter Potgieter, Monica Potgieter; Doreen Chan, Per Yhland and Peter Sham for their warm friendship, love, assistance and co-operation during my stay in Hong Kong. We spent a lot of time together, constantly helping each other. A particular note of thanks is given to Collins Yeboah Druyeh, Joseph Nsafoah Boateng and Martin Acheampong for their love and financial support.

I would like to thank my colleagues and friends: Hugo Ledoux, Ahmed Shaker, Kourosh Khosheham, Fok Hok Sum, Krzysztof Matuk, Wong Man Sing Charles, Elzbieta Matuk Carol Ng. They shared with me their research experiences and they have made my stay in Hong Kong an enjoyable experience. Many thanks are also extended to all friends and members of staff in the Department of Land Surveying and Geo-Informatics. Since the list is too long to mention, it is hoped that those who have given me help in this manner will accept this anonymous recognition.

The financial support provided by The Hong Kong Polytechnic University for this research is gratefully acknowledged. I would also like to gratefully acknowledge The Kwame Nkrumah University of Science and Technology (KNUST) for granting me study leave with pay. The continual support of KNUST was particularly significant.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

CHAPTER 1

INTRODUCTION

1.1 Introduction

In recent years, terrestrial laser scanning (henceforth, TLS) has emerged as a useful technology in a diverse range of application areas (such as heritage recording and documentation, reverse engineering, as-built surveys of industrial plants, slope stability monitoring and assessment) traditionally catered for by close-range photogrammetry (henceforth, CRP). The technology of TLS has the ability to produce high-resolution 3D surface information of an object or scene in a shorter time (i.e. near real-time) compared to classical CRP techniques. The same type of data (i.e., 3D surface information), on the other hand, can be generated using CRP techniques, but image disparities common to close-range scenes makes this an operator intensive task.

Also, TLS allows users to create 3D representation that can be visualized, measured, and modeled in many different software environments. Table 1.1 summarizes the advantages and disadvantages of CRP and TLS in terms of 3D object reconstruction. For instance, to reconstruct 3D information or to create a 3D model with CRP, more than two images must usually be taken from two different positions, the same object part must be covered and the rays to homologous points must intersect at a good angle (Kern, 2001).

Table 1.1: Advantages and disadvantages of CRP and TLS

	Close-Range Photogrammetry	Terrestrial Laser Scanning
Advantages:	High Quality RGB-texture information	High-density, near real time 3D point cloud data measurements
	Camera location is unconstrained	3D data can be visualized, measured, and modeled in many different software environments
Disadvantages:	3D model requires significant effort	Low quality RGB-texture information
		Earth bound

It then requires image matching, which includes finding the corresponding features on images, and the degree of automation in CRP is directly related to the matching problem (Schenk and Csatho, 2002). In essence, a 3D model from the CRP approach requires significant effort.

The imaging systems of some TLSs do not have very high quality RGB-texture information (i.e., lack of visual object information) whereas the high-resolution digital cameras used in modern CRP do. That is, in contrast to TLS derived points, surfaces derived by CRP are potentially rich in scene information. Also, TLSs are essentially earth-bound, whereas cameras can be moved around the object being imaged at will. This means that the camera location is unconstrained. Thus enabling a camera to capture imagery that is otherwise not possible with TLS. From these brief comparisons, laser scanning and photogrammetry can,

therefore, be seen as complementary, rather than competitive, technologies. It is possible to combine data sets from these technologies and to minimize the disadvantages of both. That is, some of the disadvantages of one method could be offset by the advantages of the other method (Schenk and Csatho, 2002).

For example, TLS can easily provide 3D surface information and a ready-made framework for CRP. CRP can readily provide high quality image texture and fill in 3D data not visible to the scanner. That is, those areas not visible to the scanner can be captured and processed using photogrammetry, and photogrammetry can make use of the scanner point cloud to define the coordinate system. Merging the CRP and TLS data could be thought of as a problem of multisensor fusion, where datasets are obtained from terrestrial 3D laser scanner and an arbitrarily located and oriented digital close-range camera.

1.2 Objectives of the Research

The prime objective of the research is to investigate the automatic fusion of high-density 3D point cloud data (such as those produced by terrestrial laser scanning systems-3D sensors) with high-resolution perspective imagery (such as that generated by digital cameras or scanned photographs - 2D sensors). This fusion should result in, amongst other things, photorealistic 3D models with the best characteristics of both; namely, high spatial resolution and high quality RGB-texture information. Much of this has been done with small-scale objects and manual processing and with the aerial imagery and surface models derived from

them, but an automatic multi-scale and multi-oriented model is a new area of work. In order to fuse a TLS point cloud and an arbitrarily located and oriented CRP image, it is necessary to register or recover the positions of the 2D cameras with respect to the 3D laser scanner (Figure 1.1).

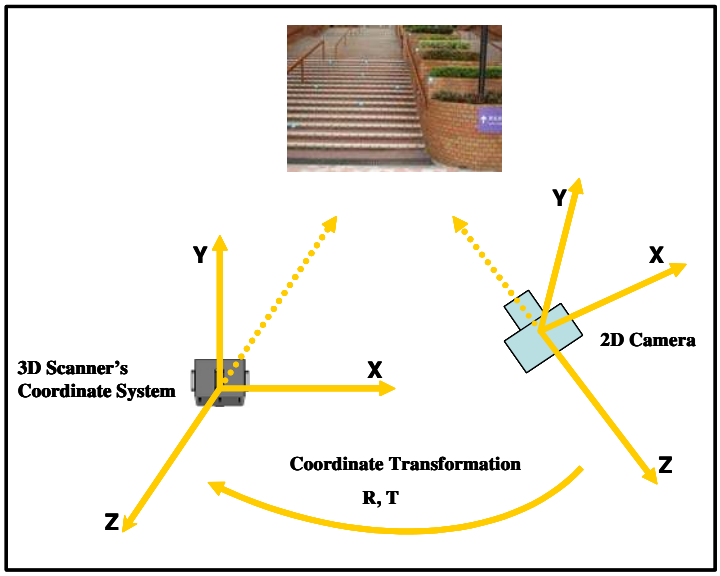


Figure 1.1: The Registration Problem

The task of this registration is to place the data (2D photographic images and 3D point clouds) into a common reference frame by estimating the transformation between the datasets (as can be seen in Figure 1.1). It is assumed that both the camera and the laser scanner view the same part of the real scene, so that the 3D and 2D views have significant overlap (Figure 1.1). Registration, which includes correspondence matching, is a prerequisite for the exploitation of the inherent advantages of multisensor systems over single sensor systems. This requires the estimation of the orientation of one sensor to the other, and it is complicated by the fact that the data sets are dissimilar (Boughorbal *et al*, 2000). That is, for

example, the camera is a high-resolution passive sensor whereas the laser is an active sensor of lower resolution.

These sensors also provide independent datasets, which differ in resolution, field of view (FOV), scale, and illumination. Although the automatic correspondence matching is not a problem for vertically oriented images, which have intrinsically much simpler geometric relationships, it is still a problem in the terrestrial case, and it is even more complex in the terrestrial multisensor case. What makes the problem more difficult for the data sets under consideration is the additional complexity of dealing with terrestrial images where orientations and locations of images are less regular. This means that locating and identifying the same object in each data set is not a trivial task (Pulli and Shapiro, 2000). It can be close to impossible due, among other things, to the difference in illumination, reflectance, perspective, and lack of appropriate texture.

Concurrent with the objective of this research is the need to:

- Develop methodology to automatically fuse 3D point cloud with close-range imagery,

- Evaluate and develop algorithms for the extraction of features from both the 2D and 3D data sets, and to assess the quality of the features; develop assessment measures for the merged data set models, and

- Development of algorithms to automate the registration process, the establishment of correspondence between objects in the 3D surface and in the 2D image, and matching.

1.3 Method

Multisensor data fusion in this thesis refers to the techniques for the combination of datasets from 3D point clouds and 2D photographic images to create a new dataset. Multisensor data fusion techniques can be performed at three different processing levels: data-level fusion; feature-level fusion, and decision-level fusion (Ma, 2001; Pohl and Genderen, 1998). This thesis addresses the problem of pixel-level and feature-level fusion using multisensor data produced by the two technologies, laser scanning and close-range photogrammetry.

Fusion at the pixel level deals with the combination of data sets from the two sensor using the photogrammetric collinearity equations whereas at feature-level sensor fusion specific information (i.e. features) from objects detected by different sensors are combined. There are two methods that could be employed in the multisensor data fusion. The first is to directly fuse the 2D image and the 3D point clouds while the second is to create a 2D representation (i.e., synthetic camera image) of the 3D point clouds. The creation of the synthetic camera image is been established as a preferable approach.

The concept and the methodology of the multisensor problem are illustrated in Figure 1.2. Figure 1.2a shows the screen capture of a high-density 3D point cloud which is displayed as a pseudo-coloured image. When the marked portion is zoomed in on, the unstructured point cloud is seen (Figure 1.2b). Figure 1.2c is 2D representation of the point cloud and Figure 1.2d is a 2D real camera image of the same scene.

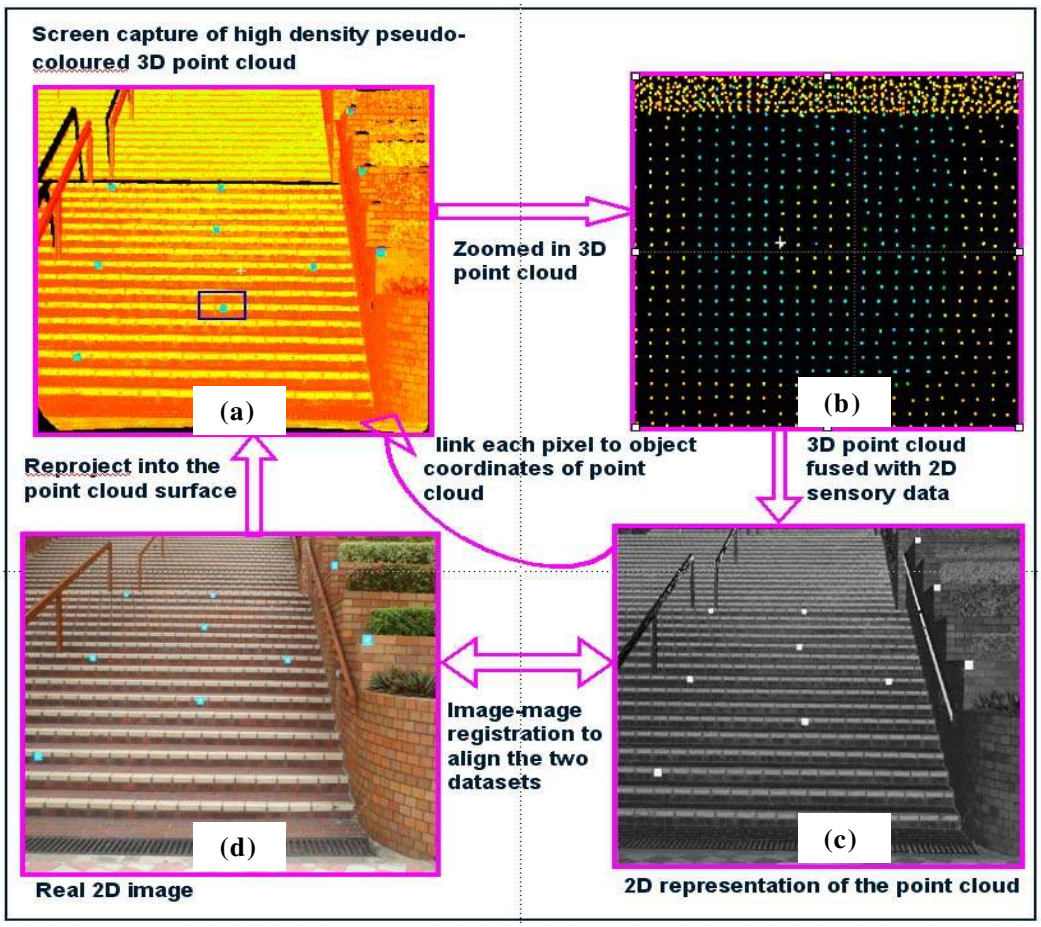


Figure 1.2: Multisensor Fusion Approach

Given these sets of high-density 3D point clouds of a scene (a) and 2D photographic imagery of the same scene (d), the questions which still need to be

answered are “can these independent complementary data sets (2D photographic image and 3D geometry) be accurately expressed in a single object-centred coordinate system; and how can each pixel in the 2D photographic image be related to its corresponding sampled 3D point on the object surface?”. As mentioned previously, the answers to these questions can be found in the concepts of multisensor fusion. This field of research has identified the following four approaches each of which are implemented here: data pre-processing (fusion of data from the two sensors), automatic measurements (feature detection and correspondence matching), mapping (creation of point cloud visual index), and orientation (calculation of exterior orientation parameters).

The first approach is the 2D representation in (b) of Figure 1.2 which is created by transforming the 3D point cloud from its raw form into a 2D synthetic camera image (SCI) and so reducing the registration process to matching this 2D image with a 2D photographic image (real camera image - RCI). This approach, which can be termed data fusion (defined as the process that combines information or parameters from the two sensors), focuses on the techniques of combining data from the sensors (digital camera and terrestrial laser scanner) to create a synthetic image. The choice of using this type of data rather than the 3D point cloud is based on the fact the SCI allows the use of already existing techniques for matching intensity images (Dias *et al*, 2001; Schenk and Csatho, 2002). That is, existing traditional image processing algorithms can operate on these generated data.

Also, to register RCI to SCI is a much easier task than to do this operation with 3D point clouds directly. This is due to the fact that the important object features (such as corner points or edges) are not directly captured by the terrestrial sensors; instead they have to be modeled from the 3D point cloud data in a separate process. In other words, the geometric features in the SCI are easier to detect than those in the 3D model. The second approach is to extract features from each image (SCI and RCI) and to determine correspondences (candidate registration points) between them for use in the image registration process. The third is to link the object coordinates of the point cloud to the pixel of the SCI. This will allow the exterior orientation of the RCIs to the point clouds.

Finally, the exterior orientation (to calculate the relative rotation and translation between the two sensors) of the images to the point clouds is computed. This research focuses on undertaking the orientation process using manual methods and existing photogrammetric application software, and automatically by using a bundle adjustment procedure. The RCIs are then mapped to the 3D geometric model to produce a geometrically correct, photorealistic view of the scene using the relative orientation and position parameters.

The fusion of close-range photogrammetry and the relatively new technology of terrestrial laser scanning can offer a significant tool in photorealistic 3D models presentation, that is texture-mapping the point cloud to create photo-realistic models which are essential for a variety of applications (such as 3D city models,

classification of real world objects, and virtual reality creation). Its other applications are the extraction of reference targets for registration and calibration purposes (El-Hakim and Beraldin, 1994); automation of the registration of point clouds, automation of 3D measurement (automatic exterior orientation); 3D reconstruction; and if the data are geo-referenced, they can be readily incorporated into existing GIS applications.

1.4 Previous work in Fusion of Multisensory Data

As previously stated, this thesis explores and attempts to provide a solution to the problems of developing a methodology to fuse terrestrial 3D point cloud data with digital close-range imagery. Most of the work done by others is with aerial imagery and point clouds from airborne laser scanning systems (ALS), which have intrinsically much simpler geometric relationships. The case at hand deals with terrestrial scanners and an arbitrarily located and oriented camera.

The review of the previous related work falls into three broad and distinct categories. The first category presents a review of general background. This presents the overview of previous related work in the area of vertically oriented image-to-image fusion, the use of laser scanner point clouds, and close-range imagery. The second category reviews work related to fusion of aerial imagery and 3D airborne LIDAR (LIght Detection And Ranging) data; and finally, previous work related to fusion of TLS and CRP methods is reviewed.

1.4.1 General Information

Previous work on image-to-image multisensor focuses mainly on the use of vertically oriented imagery (aerial photographs, satellite images, airborne laser scanners) for the small-scale representation of large objects. From the viewpoint of image-to-image fusion (fusion of only vertically oriented optical images) researchers have developed and utilized different methods to fuse complementary digital images of the same scene from multiple sensors for various applications (Sowmya and Trinder, 2000; Dare and Dowman, 2000; Yong *et al*, 2000, Saroglu *et al*, 2004, Essadiki, 2004, Sasagawa *et al*, 2004, Qi, 2004, Hong and Zhang, 2004.).

In most cases, image-to-image fusion is a preprocessing step to enhance certain features, improve classification, to substitute missing information and to detect changes in the images. The reported research showed that the use of multisensor image data has increased user's the ability to simultaneously analyze and visualize complementary information in photogrammetry and remote sensing.

Of late, considerable attention has been given to the use of digital close-range photogrammetric imagery for detailed modeling and visualization. In a series of papers, (Coorg and Teller, 1999; Pollefeys *et al*. 2001; and Chandler *et al*, 2001), various systems for extracting 3D shapes and textures from close-range image sequences are presented. They provide an initial solution to the problem of combining textures from different viewpoints in outdoor environments.

They tackled the problem of texture occlusion from the objects which are not modeled in the 3D model (such as trees, vegetation, or occluding buildings). More recently, within the photogrammetry community, several researchers have used 3D point cloud data from terrestrial laser scanners to achieve visually pleasant 3D models of outdoor and indoor objects (Kern, 2001; Barber *et al*, 2001; Thiyagarajan, 2003; Guidi *et al*, 2002; Pfeiffer and Rottensteiner, 2001).

Different modeling techniques (i.e., techniques or mapping image textures) have been applied to the point cloud to generate 3D models which are visualized in different software environments. In a series of papers (Guarnieri *et al*, 2004; Gordon *et al*, 2001a; Gordon *et al*, 2001b; Lemmens and van den Heuvel, 2001; Lichti and Harvey, 2002; Lichti *et al,* 2000; Lichti *et al,* 2002; Lichti and Gordon, 2004; Staiger, 2002; Staiger, 2003;), the quality, accuracy and errors of the terrestrial laser scanning processes are discussed. From this point of view, researchers are focusing on the use of TLS as an independent technology.

1.4.2 Fusion of Aerial Imagery and 3D Airborne LIDAR Data

In terms of fusing vertically oriented optical images and point clouds from airborne LIDAR data, a great deal has been done. For example, LIDAR, multispectral satellite image and high-resolution aerial imagery (Liang-chien *et al*, 2004), LIDAR and Satellite imagery (Antonio *et al*, 2004; Barbarella *et al*, 2004; Guo, 2003; Nakagawa and Shibasaki, 2001) and LIDAR and InSAR (Gamba *et al*, 2003), LIDAR and aerial imagery (Baltsavias *et al*, 2001; Haala and Brenner

1999; Hyung-Tae *et al*, 2000; Rottensteiner and Jansa, 2002; Schenk and Csatho, 2002; Schiewe, 2004; Vosselman, 2002;). Most of the LIDAR and aerial image fusion work has been carried out in the application area of 3D reconstruction, DEM generation and feature extraction, and the purpose of fusing the data sets has been in comparisons between photogrammetric and LIDAR derived surfaces.

An often encountered problem when using LIDAR data is the need to register LIDAR with the imagery. The quality of the registration process is a crucial factor for the fusion processing. The LIDAR data are used to supplement the existing stereo image-based extraction techniques by providing geometric constraints to guide the image matching process.

Also, the LIDAR data are used as initial 3D data to determine the search range and to detect possible occlusions. Since the accurate detection and measurement of point correspondences can be difficult, especially for the point clouds from LIDAR data, straight lines are measured between the image and the laser data as corresponding elements for the orientation process. In all the approaches, the segmentation process is used as an intermediate step to extract information on edges and linear surface features.

Other researchers like Schenk and Csatho (2002) backprojected the 3D LIDAR data into 2D by first transforming the LIDAR data into perspective projective space and then interpolating the irregularly spaced LIDAR data into a regular grid.

In addition to the advantage of visualizing the 3D LIDAR data, existing image processing algorithm can also operate on the generated image. Although the conversion was straightforward, the authors pointed out that the interpolation techniques deserved closer attention. This method is the closest to the one described in this thesis since it reduced the registration process to matching 2D data with 2D photographic data. Apart from the differences in the data sets used (i.e. aerial photographs and airborne laser scanners), they interpolated the elevation data instead of intensity values, to generate range images.

1.4.3 Fusion of Close-range Imagery and Terrestrial 3D LIDAR Data

Despite the large amount of work reported in the literature, little has been published in terms of fusion of independent close-range images with TLS point clouds. The very few that have been published in this area of research include Dias *et al* (2002), Stamos and Allen (2000), Kurazume *et al* (2002), and Elstrom *et al*, (1998).

Elstrom *et al* (1998) presented a novel stereo-based technique for registering colour and laser scanner point cloud data using externally uncalibrated sensors. The proposed registration methods took advantage of the camera images, which are a byproduct of the 3D point cloud for most commercial TLS systems from the integrated laser scanner camera. Corresponding features are automatically extracted from projective image pairs. These points are then used to estimate the relative sensor pose by minimizing the difference between scanned and predicted

stereoscopic range measurements. The sum-squared-error of the stereo and laser scanner point cloud data is then calculated, and a downhill simplex algorithm is iteratively applied to refine the rotation and translation estimates while rejecting outliers.

As outlined, although they used a Kodak DCS 460 which produced 3072 X 2060 digital images, they resampled this to a smaller size (768 X 512) for use in their registration algorithm. The requirement with this stereo approach is that there needs to be enough overlap between the TLS and the colour images and the cameras can not be placed too far apart. In addition, this method only works well in cases where the object consists of flat planes with few textures pattern and only simple line edges are extracted.

Similarly, in Kurazume *et al* (2002), the integrated camera image (Cyrax laser scanner which produces image with only 480 x 480 pixels) is used in the registration processes. This image resolution is even lower than the full scan resolution of the point clouds data (which is 1000 x 1000 for Cyrax 2500). In their method, edge features from both 3D data and 2D image are first extracted for matching. The relative position of each sensor is determined with the robust statistical method – maximum likelihood estimator - without any initial estimation of the relative sensor position. A minimum number of points is needed for the existence of robust solution. In both problems, 2D linear features are advantageous because they can be reliably extracted and are prominent in real

scenes. However, the quality of features extracted from such low resolution images is limited.

In Stamos and Allen (2000), the authors use independent 2D images and the Iterative Closest Point (ICP) approach to register the datasets. This approach requires interaction between 2D and 3D data with the subsequent problem that the corresponding sets of 3D and 2D feature are known a-priori. That is, these corresponding sets of feature matches are provided by the user. In their work, the sensors are calibrated. The authors concluded that the success of the ICP algorithm to estimate the unknown position and the orientation of the camera with respect to the range sensor depends largely on the distance between the viewpoints of the data sets and the overlap between the data sets. It can be seen that the solution to both problems requires the knowledge of a set of corresponding 3D and 2D features, that is 3D features and their 2D projections on the image plane.

None of this research has explicitly exploited the fact that the TLS and the digital camera capture their data with similar imaging geometry - perspective projection. That is, the pulses of laser light are assumed to radiate from and return to a single point inside the scanner in a fashion similar to the rays of light from the object passing through the perspective centre of the real camera lens. The conceptual elements of this problem are solved in the field of photogrammetry. Compared to the registration techniques described by previous research, the one described in

this thesis does not need any previous calibration and is automatic (i.e. minimizes user intervention). Also, the registration starts with the use of photogrammetric interest operators to find the candidate registration points. Instead of rescaling the high-resolution image, an image pyramid approach is used with the matching algorithms.

1.5 Outline of the Thesis

The thesis is organized in order to fulfill the research objectives. Chapter 2 provides the essential background to the study. The fundamental concepts of terrestrial sensors relevant to close-range photogrammetry and laser scanning are reviewed. The simplified collinearity equation used to model the multisensor data is explained in this chapter. Also, the background information on mutisensor fusion and multisensor mathematical models is presented

In Chapter 3, the methodology for the creation of synthetic camera images through multisensor data fusion is presented. The basis for establishing the synthetic data is explained, and all related information is presented. The strategy for the data capture is also discussed. This chapter also presents together the synthetic camera and real camera data used in this research.

Chapter 4 focuses on the automation of the multisensor image matching process based on the detection and extraction of interest points, and feature correspondence matching. The existing image matching strategy for feature level

fusion is reviewed. This chapter discusses the principles, concepts and implementation issues of the Harris corner detector, the zero mean normalized cross correlation and the RANSAC algorithm. The performance of the developed matching algorithm is evaluated.

Chapter 5 discusses the process of establishing a link between the pixel intensity values of the generated synthetic camera and the corresponding sampled 3D point data. The manual implementation of the orientation process used in the existing photogrammetric application software is discussed in this chapter. Also, the photogrammetric bundle adjustment with the self calibration techniques used to determine the exterior orientation parameters is discussed. This chapter discusses the process of mapping the real camera images on the 3D geometric model to produce photorealistic models of objects. Finally, the performances of both the automatic and manual methods are evaluated.

Chapter 6 compares the exterior orientation parameters as produced by photogrammetric software and bundle adjustment. Also, the exterior orientation parameters of the generated synthetic camera images are discussed. In Chapter 7, the fusion of opportunities provided by close-range photogrammetry and the relatively new technology of terrestrial laser scanning are discussed. Conclusions are drawn in Chapter 8 as an overall evaluation of this study and suggestions are made for future work.

CHAPTER 2

BACKGROUND

2.1 Introduction

This thesis develops an approach of fusing data sets from two technologies: laser scanning and close-range photogrammetry. Both are effective technologies for providing a measurement system that can record detailed 3D data. Important to this research, has been the established level of knowledge relating to the generation of the synthetic camera image, disparities of the data set and multisensor orientation, these being the concepts of multisensor fusion and the application of these to images and point clouds, and to the automation of registration processes. This chapter provides the essential background to the research. The fundamental concepts of terrestrial sensors relevant for close-range photogrammetry and laser scanning are reviewed. Also, the background information on multisensor fusion is reviewed.

2.2 Overview of Close-range Photogrammetry

Close-range photogrammetry (CRP) measures objects directly from images captured with a camera at close-range. Digital cameras, visualization and automated image measuring software, and desktop computing power, have made CRP a useful tool for reconstruction in a diverse range of applications.

Regardless of measured size, the process of photogrammetry (aerial and close-range) remains the similar, and enables the three dimensional coordinates of an object to be reconstructed from a series of two-dimensional photographs taken from various orientations around an object. To reconstruct 3D information or to create a 3D model with CRP, two images must usually be taken from two or more different positions, the same object part must be covered and the rays to homologous points must intersect with good geometry (Kern, 2001). It then requires image matching, which includes finding the corresponding features on these images.

However, in terms of the camera configuration (the locations and the orientations) of the photography and control points required, CRP has its own distinct characteristics. That is, although, close-range and aerial photogrammetry have the same basic projective geometry, the locations and orientations of the images are much less regular in close-range projects than in aerial applications (Mikhail *et al*, 2001). Camera configurations commonly used in CRP are either single or stereo camera systems, and the most popular is that of a single camera imaging an object from various exposure stations (King, 1993). In this thesis, discussion of the concepts, principles and the implementation of the CRP techniques is based on a single camera system configuration.

As pointed out in Mason (1995), high measurement accuracies can be achieved with CRP through the selection of a suitable mathematical model of the process

(for example, the bundle adjustment method), by calibration of the measurement system and with careful design of the convergent multistation camera configuration (convergent photos were used). This thesis uses a multistation camera configuration. This means that the camera axes are convergent, pointing towards the middle of the object.

The cameras that are used in CRP are metric and non-metric. Most of today's close-range photogrammetric applications (for example in heritage recording and documentation, reverse engineering, as-built surveys of industrial plants, slope stability monitoring and assessment) are carried out by means of non-metric cameras; either analogue or digital (Karras and Mavrommati, 2001). Being significantly cheaper than traditional metric photogrammetric cameras, non-metric cameras (amateur) have made analytical photogrammetry an extremely flexible and viable alternative to positioning and measurement science (King, 1993).

Measurements with non-metric cameras for CRP purposes are accompanied by: defining the image co-ordinate system (non-metric cameras do not have fiducial marks); defining the unknown elements of internal orientation (focal length and image co-ordinates of the principal point of the photograph); and defining the distortion of the lens (the distortion with non-metric cameras often amounts to considerable values and has a substantial effect).

Non-metric digital cameras use high-resolution CCD sensors in panchromatic or RGB modes, and an example of this is a Nikon D1x digital camera. This camera is used in this dissertation to capture images of the object from different viewpoints. The Nikon D1x digital camera is a 5.47 mega pixels CCD with effective pixel of 3008 x 1960. The physical sensor size is 23.7 x 15.6 mm.

The following sections provide an overview of close-range photogrammetric methods. The collinearity equations used to model the multisensor data and also used as a functional model in the bundle adjustment are reviewed in Section 2.2.2. Interior and exterior orientations models that are used for image registration and 3D reconstruction are discussed in Section 2.2.3. The photogrammetric bundle adjustment techniques used to determine the exterior orientation parameters is also reviewed in Section 2.2.4.

2.2.1 The Mathematical Model

The collinearity equations define the condition that the perspective centre PC X_o, Y_o, Z_o, a point in the object space O_1 (X, Y, Z), and its corresponding image point I_1 (x, y) must lie on a straight line at the moment of exposure (Figure 2.1). A distortion in the image signifies that there is a deviation from collinearity. Collinearity concepts are well established in photogrammetric literature and more detailed information can be found in Karara (1980).

The mathematical relationship between conjugate points on the object and image space is described by three-dimensional projective transformation which can be written in the following matrix form:

$$\begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix} = \lambda R \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix} \tag{2.1}$$

R is the orthogonal rotation matrix that describes the relative rotation of the image space coordinate axes with respect to the object space axes. The scale quantity **λ** is a scale factor of an image point.

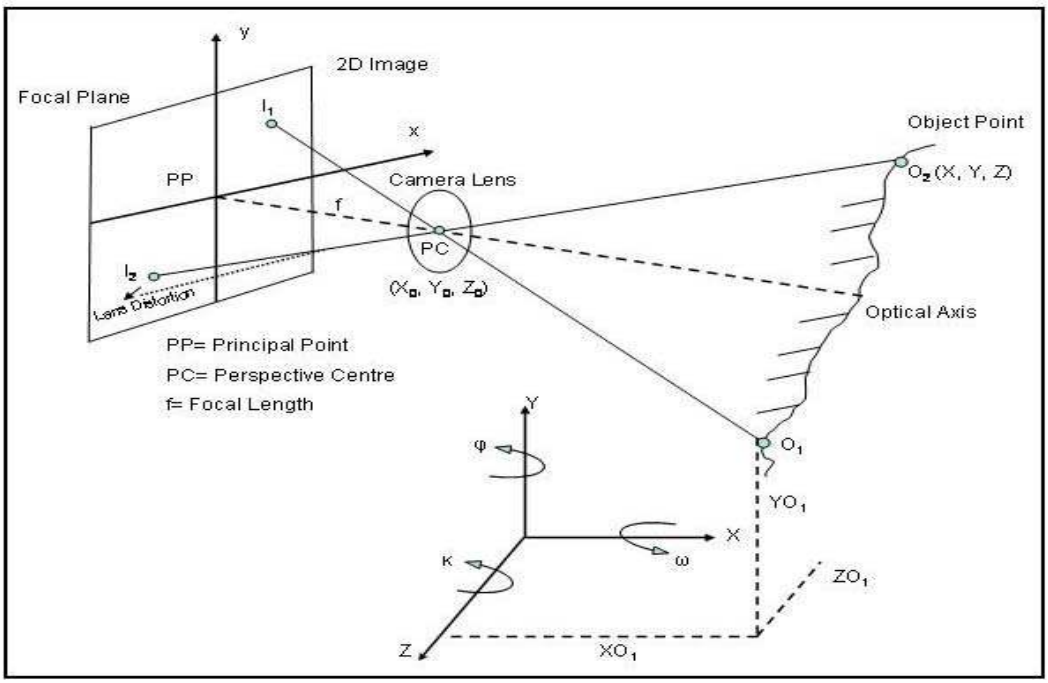


Figure 2.1: The Collinearity Model

This scale factor can be eliminated by dividing the first and second equations from the system by the third one, leading to the general collinearity in Equation 2.2 (Karara, 1980).

$$\Delta x + x - x_o = -f \cdot \frac{R_{11}(X-X_o)+R_{12}(Y-Y_o)+R_{13}(Z-Z_o)}{R_{31}(X-X_o)+R_{32}(Y-Y_o)+R_{33}(Z-Z_o)}$$

$$\Delta y + y - y_o = -f \cdot \frac{R_{21}(X-X_o)+R_{22}(Y-Y_o)+R_{23}(Z-Z_o)}{R_{31}(X-X_o)+R_{32}(Y-Y_o)+R_{33}(Z-Z_o)} \quad (2.2)$$

where:

x, y are the coordinates of image point

Δx, Δy are image distortions

X, Y, Z are the coordinates of the object space point

X_O, Y_O, Z_O are the object coordinates of the perspective centre;

x_o, y_o are image coordinates of the principal point;

R_{ij} are the elements of the rotation matrix that depend on the three angles (ω, φ, κ); and

f is the focal length of the camera. Focal length only applies when the lens is focused at the hyperfocal distance. When it is focused on a closer object, it is the principal distance, c.

f, x_o, and y_o are usually called the interior orientation parameters while $X_o, Y_o, Z_o, \omega, \varphi, \kappa$ are called exterior orientation parameters.

The implementation of the collinearity model in fusing multisensor data can be found in Section 3.3.1. Also, the Bundle Adjustment discussed in Section 2.2.3 uses this functional model.

2.2.2 Orientation Process

Orientation is the procedure where the transformation parameters from one coordinate system to a second coordinate system are determined. In photogrammetry, the re-establishment of the orientation of the camera at the moment of exposure is a necessary step in the measurement process. The processes of defining the orientation are described in this thesis as: interior and exterior orientations.

In this thesis, the orientation process is required as a pre-processing stage in the task of multisensor fusion. Existing orientation techniques fall into two broad categories: manual and automated orientation processes and the implementation of each process is described in Chapter 5.

2.2.2.1 Interior Orientation

The interior orientation determines the interior perspective of the image as it is or was at the instant of recording. Thus, it refers to the perspective geometry of the camera, and is defined by the following parameters: the calibrated focal lengths, the position of the principal point in the image plane, and the geometric distortion characteristics of the lens system.

In the interior orientation process, a geometric relationship is defined between measured or computed features in the pixel coordinate systems and image coordinate system that has its origin at the camera's perspective centre (Schenk, 1999).

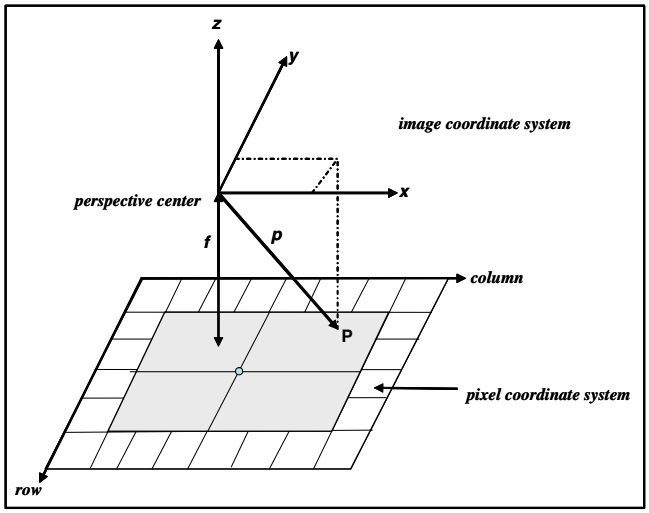


Figure 2.2: Geometry of Interior Orientation (adapted from Schenk, 1999)

The image coordinate serves as a suitable reference system for expressing spatial positions in the image space. Figure 2.2 illustrates a image coordinate systems which is commonly used to define the location of image points with respect to the exposure center. The transformation of a point ***p*** (Equations 2.3a, b and c) in the pixel system to image coordinates is (Schenk, 1999):

$$x_p = (c_p - c_o)pc \quad (2.3a)$$

$$y_p = -(r_p - r_o)pr \quad (2.3b)$$

$$z_p = -f \quad (2.3c)$$

where:

r_p, c_p are the pixel coordinates of the point ***p***;

pr, pc are the pixel size in microns;

r_o, c_o are the pixel coordinates of the principal point usually available from the camera calibration; and

x_p, y_p, z_p are the image coordinates of point p.

For digital cameras, the relationship between pixel and image coordinates is assumed constant, and is determined during the calibration procedure in addition to the parameters of interior orientation (namely the calibrated focal length, the image coordinates of the principal point and the lens distortion parameters) of the camera (Heipke, 1996; Heipke, 1997).

2.2.2.2 Exterior Orientation

The task of exterior orientation is to determine the relationship between the coordinate systems of the image and the object. There are several approaches to determining the parameters of the orientation of two or more photos. The orientation can be processed in steps (the sequential relative orientation and absolute orientation approach), with the simultaneous bundle adjustment (BA) approach, and with a Fast Fourier Transform (FFT-based) approach.

Relative orientation involves determining the relation between a pair of overlapping photos and the creation of an arbitrary 3D coordinate system (Heipke, 1997).

The transformation of the arbitrary coordinates into object coordinates is achieved in the absolute orientation phase. Absolute orientation is used to relate image space to object space after successful interior and relative orientations. The fundamental problem is to establish a relationship between image and the object space coordinates system by means of a suitable transformation.

A bundle adjustment is used compute the direct relationship between image coordinates and object coordinates. The practical procedure of absolute orientation and BA includes the search for conjugate points (the need to identify control points to be able to solve for the orientation) through image matching, the use of conjugate points to compute orientation elements, and the computation of three-dimensional coordinates of the conjugate points. In this research, the laser scanner points were used to complete this transformation process and the accuracy of orientation is determined by mainly by the accuracy of image matching and the accuracy of the ground control points (i.e., laser points).

Like the BA methods the Fourier domain approach (which use frequency domain information) can be use to register images by making use of the phase shift property of Fourier transforms. The FFT-based algorithm (discussed in Appendix A2) for automatic pairwise image registration was implemented using Matlab programming language. This thesis addresses the problem of the simultaneous BA and the FFT approaches. The FFT is evaluated in Section 5.4.1 and the BA is overviewed in Section 2.2.4.

2.2.3 Bundle Adjustment for Multisensor Measurement

An image acquired with CCD camera is, in principle, a central projection of the scene, and if more images from different viewpoints are available, the bundle adjustment (BA) method allows the computation of all camera parameters and the reconstruction of the 3D object coordinates as well as the computation of the systematic errors.

The BA employs a least squares computation based on the functional model of the collinearity equations (discussed in Section 2.2.1). There are several methods for the least squares; the method used in this thesis is based on the unified method of the least squares adjustment. Detailed information on the classical derivation of this method can be found in Slama (1980). The bundle adjustment as presented in the 3DMapper Software StereoMaker (3D Mapper, 2004) and as in SPGA program developed by Fraser (1983) was used.

The essential property of the unified method is that all variables (or parameters) involved in the mathematical formulation are considered as observations. This allows all adjustments parameters to be assigned weights that reflect their uncertainty (King, 1993). The assignment of appropriate weights (weight of an observation is inversely proportional to the observation's variance) to each parameter is critical to the performance of an adjustment. Thus, those parameter or observation whose values are considered to be well known or reliable may be assigned a high weight (King, 1993).

The total variance is the sum of the individual variances. This individual variance can be obtained by using propagation of errors (in this case error is propagated as the squared relative standard deviation). For the unified method of the least squares the value for each observation's (image coordinates, camera exterior orientation and object point coordinate observations) variance must be derived (King, 1993).

A commonly used statistic to assess the precision of parameter or observation is the standard deviation and for any measurement the error in a quantity defines how many figures are significant (Glover *et al,* 2004). That is, significant figure is the number of digits recorded in a result that reflects the precision of the measurement). The significant figures of the standard deviations presented in Chapters 5 and 6 do reflect the significant figures of the standard deviations as quoted by the manufacture of the laser scanner used in the thesis. The measurement results with the quantity zero ('0') such as in Tables 6.9 and 6.10 are denoted as '0.000'to reflect on the requisite number of significant digits.

The initial values of exterior orientation parameters are given as calculated values from the space resection (estimating the camera parameters with known 3D control points). Space intersection is also used to estimate the 3D coordinates of the object points. This requires a minimum of three control points in space whose object coordinates in the laser scanner system are known and whose image points have been measured with the imaging system to be oriented.

The accuracy of the 3D point measurements is tied to the precision of their locations (PhotoModeler, 2005), and how well they are distributed (i.e., the points are distributed evenly throughout the images) in the images. Points should be distributed over as much of the image as possible and not clustered in one corner, in order to achieve a better geometry of the object.

The BA method and its statistical model are usually extended by the incorporation of parameters for self-calibration and the use of additional parameter (AP) (King, 1993, Remondino and Borlin, 2004) functions that are intended to model the systematic image errors (i.e. the effects of radial and tangential lens distortions, or variations in lens principal distances).

Using these additional parameters, the physical process of image formation is adapted to the assumed mathematical model of central perspective represented by the collinearity equations. In other words, the additional parameters attempt to compensate for any remaining systematic inconsistencies between the mathematical model and physical reality. The procedure of self-calibration using APs introduces new observation and unknowns in the least squares estimation process, extending the bundle model and raising problems concerning the quality of the model.

2.2.3.1 Assessment of the Bundle Adjustment

The quality of the photogrammetric BA solution is partially controlled by measurement errors associated with elements used in the functional model to relate image to object. The assessment of this is primarily based on the accuracy and the precision of the computed object coordinates and, as King (1993) reports, these criteria are generally of greatest interest in close-range analytical photogrammetry. In this thesis, Root Means Square (RMS) errors of the residuals in both the object and image spaces were used to assess the accuracy of the adjustment solutions. These are determined by calculating the mean of the squares of the residuals of points from their known (true) position, and then taking the square root of the mean. The RMS errors of the object point coordinates may be computed for each of the coordinate axes (i.e. RMS_X , RMS_Y , RMS_Z) as well as for the spatial component by (RMS_P) (King, 1993):

$$RMS_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i^1 - x_i)^2} \tag{2.8}$$

$$RMS_y = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i^1 - y_i)^2} \tag{2.9}$$

$$RMS_z = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z_i^1 - z_i)^2} \tag{2.10}$$

$$RMS_p = \sqrt{\frac{1}{n}\sum_{i=1}^{n}((x_i^1 - x_i)^2 + (y_i^1 - y_i)^2 + (z_i^1 - z_i)^2)} \quad (2.11)$$

Where**:**

n = the number of points

x_i, y_i, z_i = known object point coordinates

x_i^1, y_i^1, z_i^1 = reconstructed object point coordinates

The assessment of the accuracy of each model is on the basis of RMS_P with the other RMS errors being used to indicate more trends. The statistical differences between the corresponding values of (RMS_p) from each model are analysed by the Fisher test (F-test). The F-test investigates the significance of the difference between population variances of two normal distributions, and these differences are identified via hypothesis testing. The hypothesis to be tested is the two-tailed hypothesis (King, 1993):

$H_0 : RMS_{p_1} = RMS_{p_2}$ (i.e., testing the hypothesis that the two variances are equal), versus $H_1 : RMS_{p_1} \neq RMS_{p_2}$ (i.e., larger variance (RMS_{p_1}) is significantly different from the smaller variance (RMS_{p_2}); is the F statistic being the ratio of both sample variances:

$$F = \frac{RMS_{p_1}}{RMS_{p_2}} \quad (2.12)$$

For practical reasons, the largest variance is always placed in the numerator (i.e., $RMS_{p_1} > RMS_{p_2}$). The hypothesis of equality is rejected if $F > F_{\frac{\alpha}{2},r_1,r_2}$ (i.e., F-test statistic is greater than the upper critical value), where $F_{\frac{\alpha}{2},r_1,r_2}$ is the critical value of the F-distribution with r_1 and r_2 being the respective degrees of freedom at a significance level of α. The rejection of the null hypothesis (H_0) indicates that the two RMS positional errors may come from different populations and so are considered to be different at the tested level of significance (King, 1993).

This Fisher test was used to test the statistical significance of the RMS errors of the Bundle adjustment (discussed in Section 5.4.4.1). Also, it was used to test the significance of the reference variance factor (discussed in Section 5.4.4.7) and the additional parameter estimates (discussed in Section 5.4.4.8).

The precision of the estimated parameters is expressed by the covariance matrix. Thus, the square roots of the diagonal elements of the covariance matrix give the standard deviations to the corresponding parameters which are used to evaluate the precision of the measurements system. In photogrammetric applications, the precision of the coordinates of the object points and the camera parameters can be estimated.

2.3 3D Laser Scanning Technology

3D laser scanning is a technology that captures the 3D coordinates or the digital shape of a given physical object surface in a systematic pattern (Boehler *et al*, 2003). 3D scanners may or may not deliver reflectivity values for the scanned surface elements in addition to the 3D coordinates, and they are similar, in nature, to reflectorless total stations.

As in all reflectorless laser technologies (including laser scanners), the performance is affected and limited by the physical laws of reflection, and the optical properties of materials, including refraction and inner reflection effects (Ingensand *et al*, 2003).

3D laser scanners are used as either airborne systems for topographic applications or ground-based (commonly known as terrestrial laser scanner) systems for close-range and mid-range applications. The following sections concentrate on the discussion of the 3D terrestrial laser scanner. These discussions include the principles, classification, and data processing of 3D terrestrial laser scanners.

2.3.1 3D Terrestrial Laser Scanners

Terrestrial 3D laser scanning technology provides, among other things, very dense point cloud data of objects of various sizes and shapes in near real-time. They operate at the scales of meters to hundreds of meters.

The spacing of the laser pulses (sample spacing) is adjustable to yield the required spatial resolution (i.e., number of points scanned) on a case-specific basis. That is, the resolution of the scan, which controls the number of points recorded in a scene and the level of detail visible in a scan (Barber *et al*, 2001), is simply the smallest change in angle that the scanner is capable of detecting. This resolution is also a function of sample spacing and the laser beamwidth (Lichti, 2004). Because of the radial nature of the scanner's sampling (for example, the laser beam of the Cyrax 2500 is fired from a centre point) points sampled closer will have a smaller a smaller spacing and, conversely, anything further away will have a larger sample spacing. The operator selects the appropriate resolution based on several factors such as the distance between the scanner and the target surfaces, the size of structures or process equipment being scanned and the detail required.

2.3.1.1 Classification of Terrestrial Laser Scanners

3D terrestrial laser scanners can be categorized according to their performances, such as distance or range measurement, the point density or the point accuracy, angle of view, and scanning speed.

Terrestrial laser scanners can be categorized by the principle of the distance measurement system. Mainly, the distance measurement system correlates both to the range and to the accuracy; therefore, this categorization implies a categorization by range as well as by accuracy. Depending on the principles for distance measurements, laser scanners may be categorized as: lasers using CCD

cameras where the distance measurement is based on the principle of "triangulation"; lasers using phase difference principles; and lasers using time-of-flight principles, where distance is derived from the two-way propagation time (Barber *et al*, 2001; Boehler *et al*, 2003; Lichti and Gordon, 2004).

The basic features of time-of-flight terrestrial scanners are distance measurement technologies, the beam deflection techniques which have a strong relation to maximum scanning angle, and the resolution of the encoders. These scanners use a pulsed laser beam to measure the range to a point on the scanned object (Figure 2.3). The pulse is diffusely reflected by the surface, and part of the light returns to the receiver (Boehler *et al*, 2003). The scanner measures the time it takes each laser pulse to reflect off a scanned surface and return to the scanner.

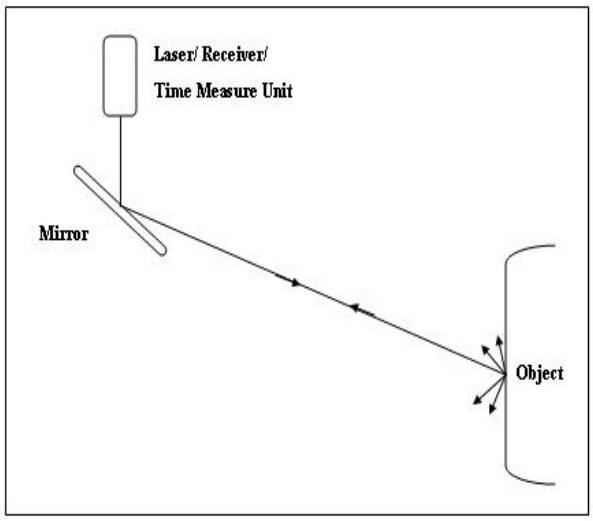


Figure 2.3: Principle of Time-of-Flight Scanner

The range, γ, is derived from the two-way flight time of the pulse, and the velocity of light, c Δt, by $\gamma = c\Delta t*1/2$. Three-dimensional measurements are obtained by encoding two rotational mirrors with orthogonal axes which deflect the laser pulse in measurable intervals of arc. In general, these scanners use a high-speed oscillating mirror system to scan the laser across a scene.

Scanners of this type include those manufactured by Cyra Technologies (Cyrax 2500 or HDS 2500, HDS 3000), Callidus Precision Systems (Callidus cp 3200, Callidus ct 180), Mensi (SOISIC) and Riegl Laser Measurement Systems (LMS-360i, LMS-420i). The collection of points from a scan is termed a point cloud that can be used directly for point-to-point measurements or for 3D modeling visualization.

A point cloud is produced with up to million points, each containing 3D geometric position (x, y, z) and a returned laser intensity (or monochromatic reflectance) value. The intensity (I), which is a relative measure of the strength of the returned laser signal, gives adequate information to distinguish points that belong to different objects or materials (Yu and Ferencz, 2001).

2.3.1.2 The Scanner used in this Study - CYRA 2500

The CYRA-System is a 4D-Laser scanner which delivers not only 3D-coordinates from the scanned points (Staiger, 2002), but also the intensity of the reflected beam. The Cyrax 2500 (recently renamed the Leica HDS2500) 3D laser scanner

consists of a scan head, and some accessories that work in conjunction with the Cyclone software that runs on a PC. Also, there are Cyra targets that are designed for use in the Cyrax system. These targets are about 15cm in diameter and the centre of the target is a reflective material of about 2mm in diameter

The popularity of Cyrax 2500 scanners is based on the accuracy at range (maximum up to hundred metres); highly adjustable scan density; high scanning speed; adjustable field of view and ease of use (Santala and Joala, 2003; Lichti and Harvey, 2002.). The Cyrax 2500 scanner is a non-contact sensor that uses a pulsed laser beam to measure the range to a point on an object's surface and uses two mirrors (Figure 2.4) to deflect the laser beam on both the horizontal and the vertical axis (Barber *et al*, 2001).

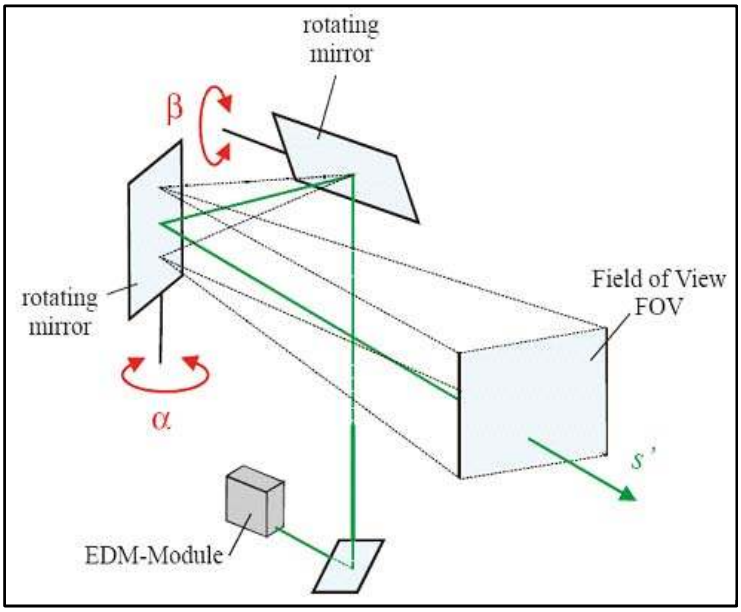


Figure 2.4: Measurement Principles of Cyra 2500 Laser Scanning System (Staiger, 2002)

The laser pulses radiate and return to a central point, thus imitating the perspective projection of a CRP camera. The rotating mirrors change the angles and emit the pulse at equal intervals of arc in both the horizontal and the vertical planes. These are the imaginary planes perpendicular to the scanner's center axis. High accuracy recording of angular settings is important, since the angles together with the distance measurements determine the reflecting point position (Boehler *et al*, 2002).

The origin of the object coordinates is "the centre of the scanner" as defined by the rotating mirrors. The XY coordinate plane is parallel to the front face of the instrument and the Z axis points to the rear of the scanner (as shown in Figure 2.5).

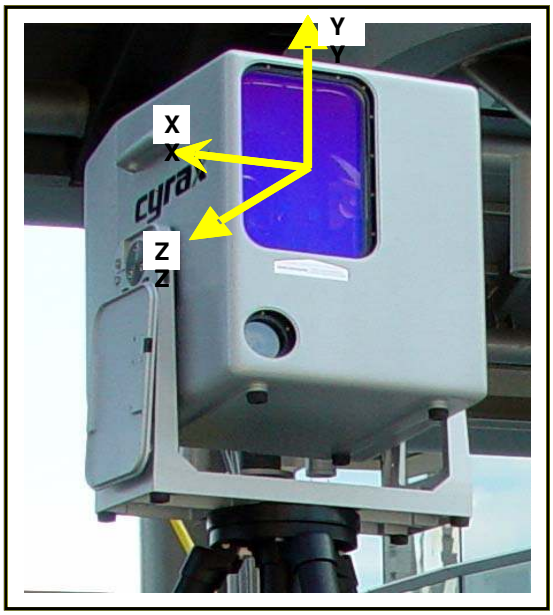


Figure 2.5: Cyrax 2500 Laser Scanner Coordinate System

The same coordinate system definition is used in CRP, the only difference is that in a CRP, there is an image plane onto which rays from an object and passing through the perspective centre are projected. This means that the TLS systems (example Cyrax 2500) and a digital camera capture their data with similar imaging geometry. This fact can be exploited in that it allows the 3D point cloud to be transformed into a 2D image (discussed in Section 3.3.1) by using the collinearity equations (discussed in Section 2.2.2).

As mentioned before, the Cyrax 2500 uses a reflectorless technology and, therefore, the performance is affected and limited by the physical laws of reflection of the laser beam by the object surface, the optical properties of materials, and refraction. One of the salient properties of a scanner that strongly influences both point cloud resolution and positional uncertainty is the laser beamwidth (Lichti and Gordon, 2004). Every point cloud produced by a scanner contains a considerable number of points that show gross errors (Boehler *et al*, 2003).

In this thesis, the calibration of the Cyrax 2500 is not discussed. That is, the quality of the measurements, such as the accuracy of the 3D coordinates, the angular accuracy, and range accuracy were not investigated.

Table 2.1: Technical Data of Cyrax 2500

Technical Data	Performance
Single Point Accuracy	
- position	±6mm (1σ) @ 1.5 - 50m range,
- distance	±4mm (1σ) ,
- angle	60micro-radians (12")
Modeled Surface Precision	±2mm
Beam Spot Size	6mm from 0 – 50m
Maximum Field of View	40˚ vertical and 40˚ horizontal
Scan Rate	1000 points/sec
No. of Measured Points	1000 x 1000 points/scan
Range	maximum up to 100m
Laser Class	Class 2 pulsed green laser

However, the technical data (which can be accessed at http://hds.leica-geosystems.com/products)) as listed in Table 2.1, has been reported and evaluated by Boehler *et al* (2003), Lichti *et al* (2000), Lichti and Harvey (2002), and Santala and Joala (2003). The reported research showed that the Cyrax 2500 laser scanner (used in this research) fulfills the specifications given by producer. These specifications were used (in Chapters 5 and 6) as the benchmark against which the accuracy (in terms of RMS errors of the residuals) of both manual and automatic bundle adjustment solutions were compared.

2.3.1.3 Measurement Errors and the Nominal Accuracy of the Cyrax 2500 Laser Scanner

There are three types of measurement errors (gross errors, systematic errors and random errors) that are likely to affect the quality of the photogrammetric BA solutions. The size of each of these errors in each measurement affects the accuracy and precision of the computed parameters (these are the positions and rotations of the camera and the image and the objects coordinates of points on the object).

The nominal accuracy of the Cyrax 2500 laser scanner (used in this research) was used as the benchmark against which the accuracy (in terms of RMS errors of the residuals) of both manual and automatic bundle solutions were compared. The nominal accuracy (used to assess the performance of the bundle solutions) for a single point are: range - Se_z = ±4 mm, and position in the XY plane (Se_{xy}) = ±6 mm.

As previously pointed out, these standard deviations of the Cyrax 2500 laser scanner are quoted for one sigma (1σ). This means that there is a 68% probability (i.e., 68% confidence interval) that the "true" mean value of any measurements fall between +1 and -1 standard deviation. Since this is not a very strong probability, the performance of each method for each data set was assessed based on the 95% confidence interval (i.e., 95% confident that the "true" value lies between -1.96 and +1.96 times the standard deviation).

That is the comparison simply compares the measurement errors obtained by each BA to these nominal values for each data set. Any measurement error greater than two times the standard deviations (2 Se_{xy}, and 2 Se_z) of the nominal accuracy (for both position and range) of the scanner was considered to be an error (gross error, systematic error and large random error) and it therefore, requires further investigations for a more conclusive evaluation.

2.3.2 3D Terrestrial Laser Scanner Data Processing

The result of the laser survey is a very dense cloud of points (also called a point cloud). For each point of the model, the X, Y, and Z coordinates and the reflectivity value are given. This cloud of points can be used to describe and model the scanned object. Point clouds can be easily visualized in three dimensions and the necessary surface, line or point data can be extracted using either Cyra's Cyclone software or CAD systems such as AUTOCAD and Microstation. Important object features, such as corner points or edges, are not directly captured; instead they have to be modeled from the point clouds in a separate process (Boehler *et al*, 2003).

Processing point clouds (the conversion from the discrete points into surfaces and solid modeling) requires that an operator to interact with the points and be able to understand what the point cloud represents. That is, the laser scanner data processing consists of a set of actions (such as the fitting of geometric and mathematical surfaces) that are necessary to obtain the correct digital model of the

object. The operations that are directly carried out on the point cloud include, for example, data filtering (noise reduction), point cloud registration (combining several scans with different orientations and positions) and georeferencing (the orientation of point cloud to an absolute coordinate system). The Cyclone software allows camera images to be overlaid on the surface model, but, as already pointed out in Section 1.4.3, the resolution of these images from the laser-based cameras is typically low.

2.4 Overview of Multisensor Data Fusion

While the concept of image fusion is not new, the emergence of new imaging sensors has created a need for image processing techniques that can fuse observations from a variety of different sensors. In image processing or in remote sensing applications, image fusion (multisensor, multi-temporal and multi-resolution imagery) is commonly undertaken, using a variety of techniques. There is also an increasing number of studies that combine information on different types (e.g. imagery with ground elevation data or magnetic readings). For instance, in this thesis, multisensor fusion combines the terrestrial 3D point cloud and 2D photographic imagery. Although many papers focus on the topic of 2D image fusion, there is little research on the topic of 3D data fusion.

Multisensor fusion is becoming a standard source of information for various mapping applications, including orthoimage production, feature extraction,

surface reconstruction and mapping, perspective view generation, and virtual reality (fly-through) creation.

In image fusion, the data may differ in sensor type, viewing conditions, camera position, or capture time. Depending on the synergy of the information inherent in the data, it is possible to enhance the spatial, temporal, and or spectral resolution, extract the 3D structure, and improve the decision-making (detection, classification, recognition, etc.) performance (Gunturk, 2003).

In the following sections, a general introduction (which includes the concept and the techniques) of multisensor data fusion is provided. The description of levels of multisensor data fusion, problems and issues of multisensor data fusion, and multisensor data fusion algorithms are presented. The complexities involved in the multisensor data sets used in this thesis are also discussed.

2.4.1 Concepts of Multisensor Data Fusion

Lee *et al,* (2002*)* describe multisensor data fusion as the process that combines data and information from multiple sensors (typically provided by several images or aspects of a scene), obtained using different imaging techniques or with different acquisition parameters or at different times, in order to achieve improved accuracy and better inference about the environment than can be achieved by the use of a single sensor. The fusion of multisensory data provides complementary information, as sensors based on different physical principles register different

properties of the object. This, in turn, translates to a more consistent scene description enabling an improved scene interpretation and understanding.

Applications of multisensor fusion may be characterized by the level of representation given to data during the fusion process, and by the mathematical foundation upon which the fusion operation is based. The goal of multisensor fusion is to create new images that are more suitable for the purposes of human visual perception, object detection and target recognition (Li *et al*, 1995), than those from any of the individual sensors. However, such sensors provide independent measurements with a very large spectrum of characteristics with respect to data rates (spatial and temporal sampling), spatial and temporal scales, field of view and means of acquisition (Csatho and Schenk, 1998). This diversity creates a tremendous source of practical problems, whose resolutions lie in a good understanding and modeling of more fundamental components (Li *et al*, 1995).

2.4.1.1 Complexities in the Data Sets

Fusion of multisensor data sets is a complex task, as it requires the application of mathematical concepts and algorithms. The choice of a technique or algorithm for combining data in a multisensor system depends mostly on image (Vajdic *et al*, 1995) dimensionality (2D, 3D), image types (natural scenes, stereo images, satellite images or medical images), transformation (rigid, affine, projective, curved), search determination (direct, search- oriented), and interaction (semi-automatic, automatic).

A prerequisite for successful multisensor data fusion is that multisensor data have to be correctly registered or relatively oriented. Thus, one of the processing steps is to perform a multisensor registration or data alignment. Since each sensor provides data in its local frame, data from different sensors must be converted into a common reference frame before combination. This requires the estimation of the orientation of one sensor to the other (see Chapter 5), and it is complicated by the fact that data sets are dissimilar. That is, for example, the camera is a high-resolution passive sensor whereas the laser is an active sensor of lower resolution.

The second key element of the sensor fusion task is the occlusion or obstruction problem. Since the sensors are offset by some translation and rotation, it is possible that certain segments of the scene might be occluded for one sensor while not for the other. Even a small offset may lead to occlusion, and resolving the occlusion is important to establishing proper correspondence among the points in the dissimilar data sets. The correspondence problem is a difficult one as the feature appearance can change in different views, and other similar features may exist.

Another important issue is the dynamic range between the two images. Since different sensors provide these images (discussed in section 3.3.3), the dynamic range of the two images must be re-sampled to match one another. For example, the camera is a high-resolution passive sensor in panchromatic or RGB colour mode, whereas the TLS is an active sensor of lower resolution in monochromatic

reflectance. These sensors also provide independent datasets, which differ in resolution, field of view (FOV), scale, illumination and reflectance. Table 2.2 summarizes some of the disparities between the data sets. Also, Point clouds represent a metaphor of "layers" which are transparent, so that offset parts or regions between two overlaid clouds in different colours can be identified (Naai-Jung, *et al,* 2004) whereas photographic imagery can be generally described as opaque (no light is transmitted through it).

Table 2.2: Disparities in the Data Sets between TLS and CRP

	Point Clouds/TLS	**Photos/CRP**
Resolution	1000 x 1000	4000 x 3000 pixels
Sensor Field of View	40˚x 40˚	40˚x27˚
Sensor Dynamic Range	Monochromatic Reflectance (10 bits)	RGB (8 bits per Channel)
Data Opacity	Transparent	Opaque

2.4.2 Classification of Multisensor Fusion Techniques

There are several criteria to categorize current multisensor fusion techniques. These criteria include types of sensor data, mathematical multisensor fusion algorithms and levels of representation (Ma, 2001). The following discussions are based on the levels of representation.

2.4.3 Levels of Representation of Multisensor Data Fusion

Multisensor data fusion techniques can be performed at three different processing levels: (a) data-level fusion; (b) feature-level fusion, and (c) decision-level fusion (Ma, 2001; Pohl and Genderen, 1998).

This thesis addresses the problem of data-level and feature-level fusion using multisensor data produced by close-range laser scanners and digital cameras.

2.4.3.1 Sensor Data Level Multisensor Fusion

Sensor data-level fusion (that is fusing data at the sensor level) is, in general, only feasible between identical sensory devices (Bruder *et al*, 1994). Data-based fusion occurs when the raw data from various sensors are combined without significant post-processing (Nickels *et al*, 2003). The most common techniques for sensor data fusion are grouped into: colour techniques (RGB, IHS), statistical approaches (such as Principal Component Analysis and Regression Variable Substitution), numerical techniques (averaging), neural networks, wavelets, fuzzy logic, contrast pyramid and functional models (Schiewe, 1999). In this thesis, functional model based on photogrammetric collinearity equations are used to fuse data from the multisensors (see Chapter 3). That is, at the data level, the physical parameters measured by the sensors are combined.

2.4.3.2 Feature- Level Multisensor Fusion

Feature-level sensor fusion is the process whereby specific information (i.e. features) from objects detected by different sensors are combined (Nickels *et al*, 2003). Using features to represent the sensory information not only reduces the complex of the processing procedure but also increases the reliability of the processing results (Ma, 2001). Feature-level fusion can be divided into feature extraction, feature correspondence, and registration. Typical features extracted from an image and used for fusion include edges and regions of similar intensity (discussed in Chapter 4).

Feature based methods have been shown to be more suitable for the problems of multisensor image registration (Hsieh *et al*, 1997). Images are represented by a set of features after extraction, either in the spatial domain or in the transform domain. Different techniques have been developed to perform feature-level fusion. These include techniques based on wavelet domain, expert systems neural networks, Bayesian inference and Dempster-Shafer reasoning.

Establishing correspondence between projections of the same feature in different views is performed based on finding image characteristics that are similar in both views. A vast number of corresponding algorithms have been proposed. These include epipolar constraints, uniqueness constraints, photometric compatibility constraints, and geometric similarity constraints. The geometric constraint is based on the observation that geometric characteristics of features (such pixels on

edges, lines and corners etc.) found in the first and the second images do not differ greatly.

Developing a robust algorithm to establish correspondence is one of the most important tasks in automated image registration. In most existing feature-based techniques, feature correspondence is still the most challenging problem. Features are extracted from each of the sensor data, followed by a registration step.

This requires robust feature detection and selection of an algorithm capable of recognizing features in multisensor images (see Chapter 4). The idea is to identify some features (edges, corners, lines etc) on the images from each separate sensor and then to combine these features in an overall database. In this thesis, the multisensor feature detection technique is based on the Harris corner detector (Harris and Stephens, 1998).

2.5 Summary

The Chapter provided the essential background to the research. An overview of close-range photogrammetric methods was presented. This included the fundamental principles and concepts of CRP. The collinearity equations used to model the multisensor data and used in the bundle adjustment were reviewed. Photogrammetric interior and exterior orientations were described. The photogrammetric bundle adjustment technique used to determine the exterior orientation parameters was also reviewed.

The fundamental concepts of terrestrial sensors relevant for laser scanning were also reviewed. These discussions include the principles, classification, and data processing of 3D terrestrial laser scanners. The Cyrax 2500 (the Scanner used in this study) scanner was described.

Also, the background information on multisensor fusion was reviewed. A general introduction which included the concept and the techniques of multisensor data fusion was provided.

The description of levels of multisensor data fusion, problems and issues of multisensor data fusion, and multisensor data fusion algorithms were presented. The complexities involved in the multisensor data sets used in this thesis were also discussed.

CHAPTER 3

CREATION OF SYNTHETIC IMAGE THROUGH MULTISENSOR DATA FUSION: DATA LEVEL FUSION

3.1 Introduction

In this context, multisensor data fusion refers to the techniques for the combination of datasets from 3D point clouds and 2D photographic images (i.e. the intrinsic parameters of the CCD camera) to create a new dataset. These sensors are not calibrated. Given these sets of a high-density 3D point cloud of a scene and 2D photographic imagery of the same scene, the question which still needs to be answered is "can these independent and complementary data sets (2D photographic image and 3D geometry) be accurately expressed in a single object-centred coordinate system?"

This section attempts to answer this question by creating a 2D representation of high density 3D point clouds. This is achieved by transforming the 3D point cloud from its raw form into a 2D synthetic camera image (SCI) and so that the registration process is reduced to matching this 2D image with a 2D photographic image (real camera image - RCI). As mentioned before, the choice of SCI rather than the 3D point cloud is based on the fact that the SCI allows the use of existing image processing algorithms.

Another factor which influences the use of the SCI is the fact it easier to detect and identify important object features (such as corner points or edges) in SCI than in 3D point clouds. These object features are not directly captured by laser scanners; instead they have to be modeled from the 3D point cloud data in a separate process.

In the following sections, the process of creating a synthetic image through multisensor data level fusion is discussed, and all related information is presented. The strategy for the data capture is also explained. The synthetic and real camera images for different scenes are presented

3.2 Data Set used in the Study

In order to assess the fusion of 2D images and 3D point clouds, an experiment was conducted. The fieldwork was carried out on the University Campus where different data sets with geometric structures were selected. Focusing on the amount of geometric primitive, the data sets are grouped into three scenes (data set A, data set B, and data set C) representing three levels of complexity (low, medium and high). Each data set consists of one SCI and four RCIs. A low complexity scene (data set A) contains simple continuous geometric features (figure 3.2 (a) is an example). Figures 3.2 (b) and 3.2 (c) are examples of objects that represent medium (data set B) and high (data set C) complexity scenes, respectively, and contain many geometric features of different sizes and shapes.

To facilitate terrestrial measurements and comparison between the photogrammetric and laser scanning systems, retro-reflective Cyra-targets compatible for both sensors, were placed in data sets A and B (see Figures 3.2a and 3.2b). These Two data sets are used in the manual process with photogrammetric software. The results of these two data sets are presented and discussed in Chapters 5 and 6. All three data sets (A, B, and C) are used in the automatic process with a photogrammetric bundle adjustment. The results of these data sets are presented and discussed in Chapters 5 and 6.

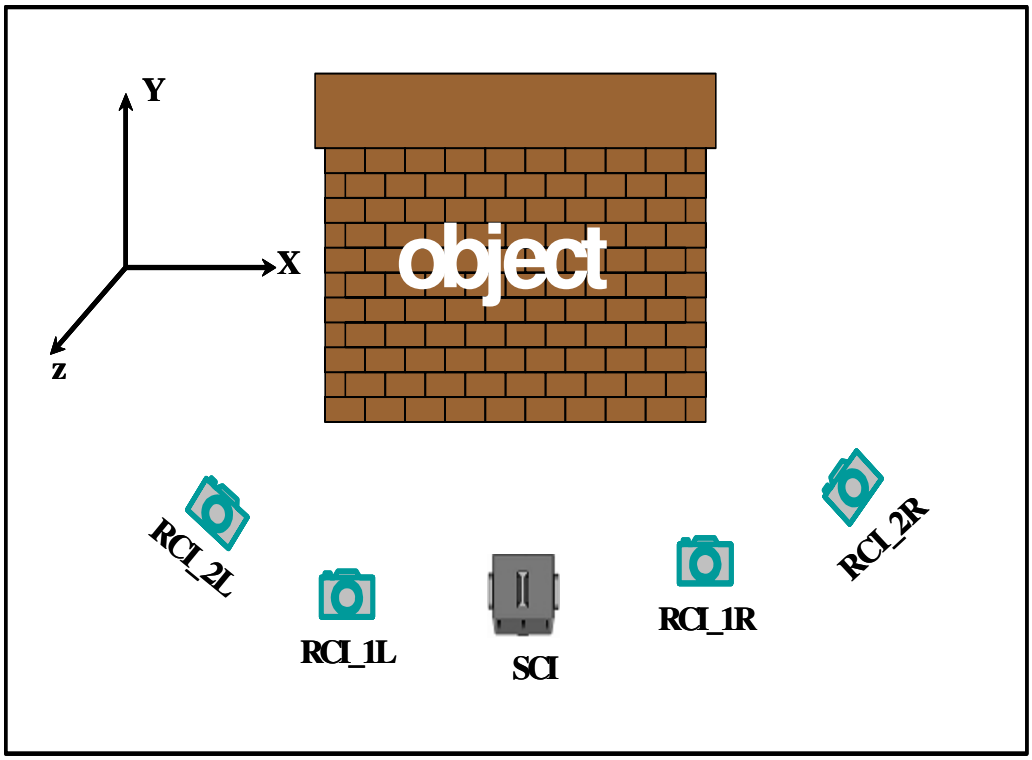


Figure 3.1: Data Capture Setup

Figure 3.1 shows the relationship between the data acquisition systems, where the laser scanner is positioned in the center (SCI) and that of the digital camera to the right (RCI_1R and RCI_2R) and then to the left (RCI_1L and RCI_2L) of the scanner. The idea is to form a multisenor model (each model is composed of two images) of SCI with the 2D images from various camera positions (or viewing angles). In particular, this methodology focuses on a single point cloud with multiple images. To test both the registration and the correspondence algorithms, three scenes were scanned with the Cyrax 2500 laser scanner. For each scan, a 2D photographic image was taken with a Nikon D1x digital camera equipped with a 20mm lens.

3.2.1 2D Real Camera Images

A series of images was taken in different directions and positions (as depicted in Figure 3.2) by the digital CCD camera (Nikon D1x). These images are called real camera images (RCIs), and one (out of four for each data set) is represented in Figure 3.3. That is, Figure 3.2(a) is an example of the RCI for data set A, Figure 3.2 (b) for data set B and Figure 3.3(c) for data set C. The remaining three RCIs of each dataset are shown in Appendix A1. This camera provides a digital image with a resolution of 3008 by 1960 pixels at true colour mode. The pixel intensity depends on the surface reflectivity parameters, surface orientation, type and position of illuminants and the position of the viewer.

Figure 3.2: Example of Real Camera 2D Images.

3.2.2 3D Point Cloud Data

A Cyrax 2500 Laser Scanner was used to carry out the laser scanning to acquire a discrete representation of the object. A point cloud was produced with one million points, each containing 3D geometric vectors (x, y, z) and a returned laser intensity-colour information-value.

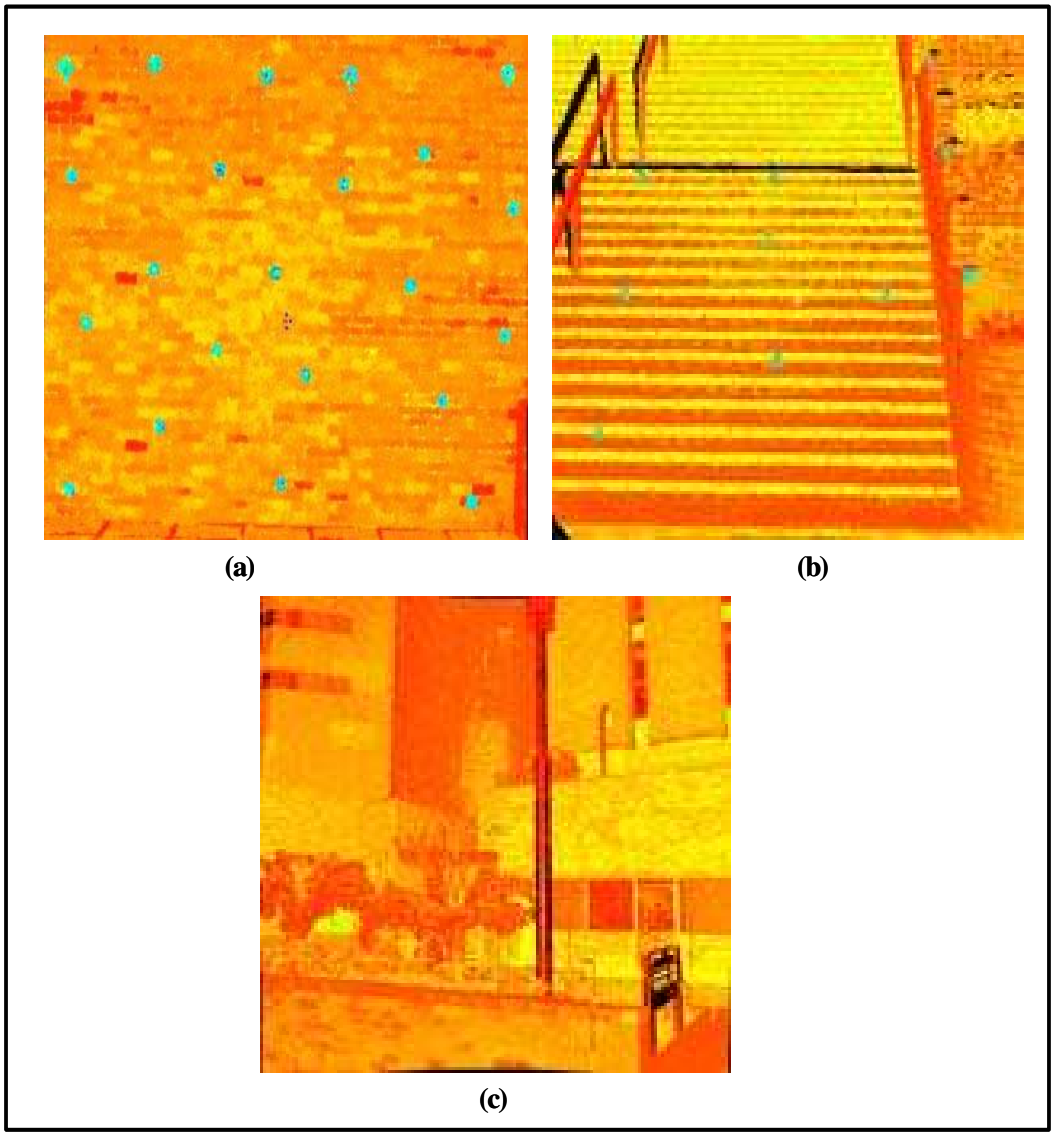


Figure 3.3: Examples of Screen Capture of high-density Pseudo-coloured 3D Point Clouds of the Three Data Sets

These coordinates can be stored as ASCII formats for input into nearly any CAD programe or point-cloud processing software. Screen captures of 3D point clouds of three test areas as displayed in Cyra's Cyclone Software are shown in Figure 3.3. Figures 3.3 (a), (b) and (c) are the screen capture of pseudo-coloured 3D point clouds for the three sets of data (i.e., data sets A, B and C, respectively).

Table 3.1 summarizes the point cloud information about each data set. All data sets were scanned at full field of view (40° horizontal x 40° vertical) and at full resolution. The 3D point cloud allows for the construction of a 3D surface model of the scene. Cyra targets were distributed through the scan and were used as control points for the orientation. These were scanned at a high density to ensure accurate geometric centers.

Table 3.1: Summary of Information of Point Cloud

Data Set	**A**	**B**	**C**
Nominal Spacing (m)	0.004 x 0.004	0.006 x 0.006	0.004 x 0.004
Max Range(m)	5.6	21.2	43.0
Closest Point(m)	1.3	4.5	3.4
No. of Points	999548	995564	973862

3.3 Creation of Synthetic Camera Image-SCI

One characteristic of TLS systems that makes them similar to a camera is that they capture their data with similar imaging geometry – that of a perspective projection. This fact can be exploited in that it allows the 3D point cloud to be transformed into a 2D image (i.e. creating a synthetic image equivalent to that of a digital camera) and so the registration process is reduced to matching these 2D data with a 2D photographic image.

Once created, the synthetic camera image (SCI) can be used for many tasks including, but not limited to, the exterior orientation of independent 2D photographic images (real camera images-RCIs) to the 3D terrestrial point clouds. This will then allow texture mapping and data augmentation through the application of close-range photogrammetric techniques, and the registration of the of overlapping point clouds. These applications utilize the two dimensional properties of the SCI instead of three dimensional interaction with the point cloud.

The following sections outline the methods used to fuse information between the two sensors (a real camera and terrestrial a laser scanner). The two sub-tasks contained in this data integration procedure include the transformation of 3D points into 2D synthetic intensity images and the computation of 2D mages (x, y) from the 3D point cloud data (X, Y, Z) as produced by a 3D laser scanner. The former is discussed in Section 3.3.1 and the latter is discussed in Section 3.3.2.

3.3.1 Strategy for Creating the Synthetic Data

This section describes the fusion of the information provided by the digital camera and the laser sensors to create a synthetic data source. Data fusion, in this section, refers to the fusion of information between the two sensors. This requires knowledge of the internal camera parameters (assumed principal distance, principal point location and no lens distortion correction). The knowledge of these parameters facilitates the creation of the synthetic data using the collinearity equations discussed in Section 2.2.2.

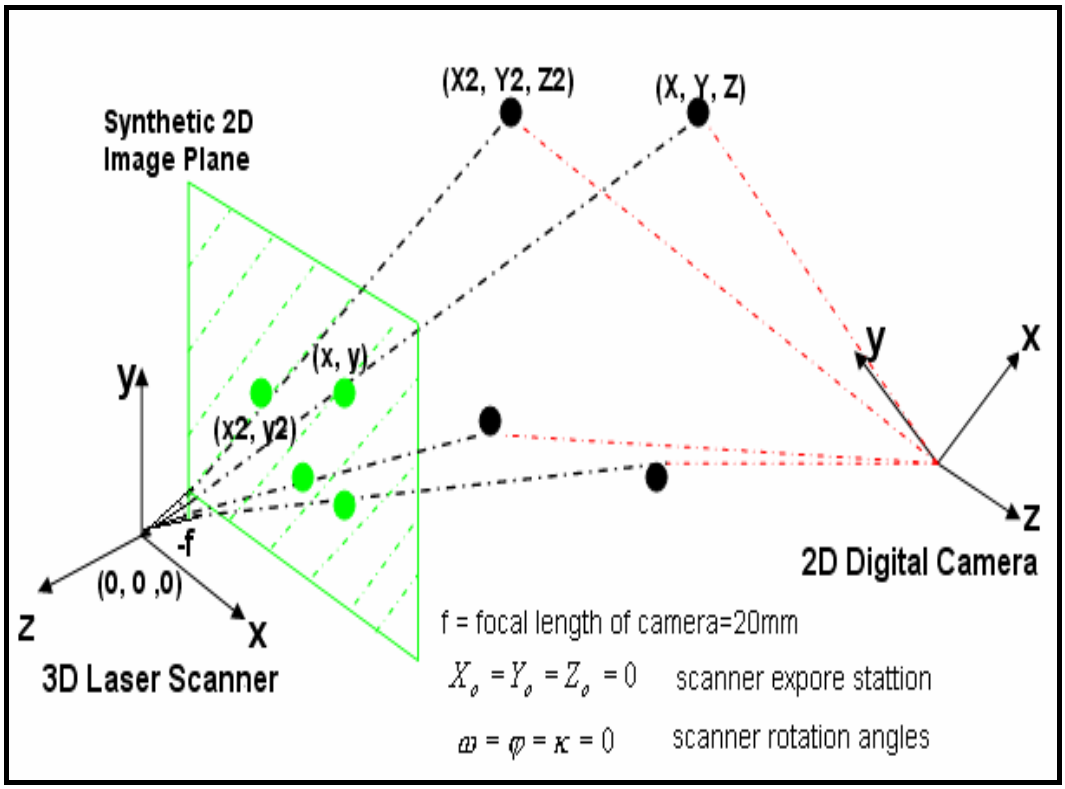


Figure 3.4: Collinearity Model for Scanner and Camera

In this research, the integration setup consisted of 3D data from a 3D laser scanner and 2D photographic images from an arbitrarily located and oriented CCD camera. As previously pointed out, the two sensors had the same coordinate definition, the only difference was that in the CRP, there was an image plane onto which rays from object and passing through the perspective centre were projected.

This positive image plane (usually used in CRP) is parallel to the xy plane and lies along the negative z at a distance equal to the focal length of the camera. If a similar image plane is placed in front of the laser scanner (as can be seen in Figure 3.4) at a distance f, (the focal length of the real camera) and the point cloud data is backprojetced on it, an image of the points equal to that taken by the camera with

the same principal distance can be generated. The relationship between a point cloud, TLS coordinate system, the SCI coordinate system and the backprojected image is illustrated in Figure 3.4.

With a known 3D surface (as produced by a 3D laser scanner in X, Y, Z), the coordinates of the exposure center (X_0, Y_0, Z_0) of the laser scanner, the rotation angles ($\omega=\varphi=\kappa=0$) the focal length of the laser scanner (assumed to be equal to that of the digital camera - f = 20mm) and basic knowledge about CCD camera characteristics (such as the interior orientation parameters), and the position of each point of the point cloud on the synthetic image are given by:

$$x = -f\left(\frac{X}{Z}\right) \qquad (3.1)$$
$$y = -f\left(\frac{Y}{Z}\right)$$

where X, Y and Z are the object coordinate of the points in the point cloud and x and y are the image coordinates of the SCI. For given a point in the laser scanner object space (X, Y, Z), its corresponding point in the 2D image plane (x, y) was estimated. The collinearity equations were implemented using Matlab programming language. Thus, the synthetic data formation, in principle, is described by mathematical rules of central perspective formation of the synthetic image.

3.3.2 A Method for Transforming Synthetic Data into Synthetic Images

Although an even point spacing during scanning was specified (as can be seen in Table 3.1) in both the horizontal (X) and vertical (Y) directions at a specified scene range, upon back projection it was found that the points exhibited some variation in spacing (as can be seen in Figure 3.5). This is due to the fact that point spacing generally depends on the shape of the object and the variation in the range from the scanner.

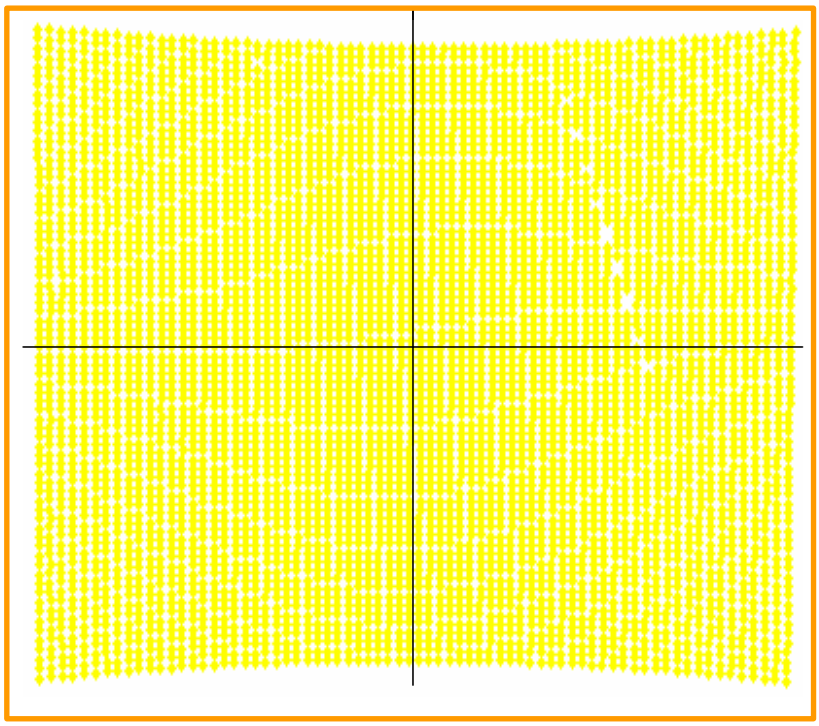


Figure 3.5: Example of Backprojected Point Cloud

To create a synthetic image equivalent to that of a digital camera, the backprojected points (discussed in the previous section) must be interpolated onto a regular grid (the pixels of the images) at an even spacing.

The purpose of this interpolation is to transform the random set of point in (x, y, I) into a uniformly spaced rectangular grid. In other words, interpolation takes a series of (x, y, I) points and generates estimated values for I (intensity) at new (x, y) points.

Resampling has two components, namely the resampled locations and the resampled values (Morgan and Habib, 2002). The former has to deal with the pixel size or the grid intervals when we want to resample into a regular grid, while the latter deals with the interpolation function. The difference among various interpolation approaches lies in the interpolation function that is chosen. Many of these are discussed in Dakowicz and Gold (2003).

The size of the grid spacing has a strong influence on the errors introduced during interpolation (Smith et al, 2003b). Errors are always present in the interpolated data, but the complexity of the interpolation method used will depend on the difference in the density between the original point cloud and the pixels in the synthetic image. Behan (2002) and Smith *el al* (2003b) quantified errors within models produced from different interpolation algorithms. They concluded that the optimal spacing should be as close as possible to the original point spacing. That is, the most accurate models were created using grids which had a similar spacing to the original points.

Based on a high scanning density cloud of 3D points (with a maximum point spacing of six millimeters) the nearest neighbour interpolation (where the value at the interpolating point is the same as the value at the grid point closest to it) was used in this thesis. Thus, nearest neighbour interpolation determines the intensity value from the closest pixel to the specified input coordinates, and assigns that value to the output coordinates. As well as being the most efficient method in terms of processing or computation time, the nearest neighbour approach does not alter the original intensity values that are used to create the SCIs (Fritscher, 2004). That is the intensity values in the SCIs remained unchanged and this is particularly important when dealing with pseudo-coloured intensity values where the palette of colours is already limited. Preservation of the original intensity values can also be an advantage since SCIs will be used in automatic measurements such as feature extraction and correspondence matching (discussed in Chapter 4).

However, nearest neighbour interpolation may result in some intensity values being duplicated while others are lost. As previously pointed out, the grid spacing (8mm) was as close as possible to the original point spacing (6mm) and as can be seen in Figures 3.8, 3.8 and 3.10, the out images (SCIs) do not appear jagged or blocky. Also, the fidelity of the concept of generating synthetic camera images has been tested by determining the exterior orientation of the synthetic camera images and the real camera images relative to the point cloud.

There are two options related to this interpolation. The first option is to generate the SCI by keeping the original resolution of the point cloud data and to then compute a new pixel size. This means that the resultant image will always have 1000 by 1000 as row and columns, and the pixel size will be as close as possible to that which fits the backprojected point clouds.

The second option, on the other hand, is to keep the pixel size of the real camera image (the Nikon D1x has 8μm pixel size) and to then compute the number of pixels. To create a synthetic image equivalent to that of a digital camera, the second option was used to generate the synthetic images shown in Figures 3.6, 3.7 and 3.8. In this case, the generated SCIs and the RCIs will have the same pixel size and the resolutions of the SCIs may be as close as possible to that of the RCIs. The density of the generated grid is a function of the minimum, maximum and the increment values for both x and y. Once the pixel spacing is defined, the interior orientation of the SCI is also defined.

3.3 3 Dynamic-Range Compression

Dynamic range refers to the range in which the scanner is cable of measuring the intensity. Grey scale images consist of pixels, each of which indicates a level of intensity. For an image of eight bit integers, the shade is in the range [0; 255]. As can be seen in Table 3.2, the generated synthetic images of all the three data sets are high dynamic range images. That is the range of intensities is larger than the standard 8-bit-per-channel (256-level) range. The reflectance values as produced

by the Cyrax 2500 for all the data sets is in the range [1199, 1256]. This means that some shades may be difficult to distinguish from each other because they are so close together that human vision cannot distinguish them from each other.

Table 3.2: Dynamic Range Values

Data Set	Minimum & Maximum Intensity Values	Dynamic Range Values
A	[-429, 770]	1199
B	[-479, 771]	1250
C	[-482, 774]	1256

The ability to distinguish shades of grey in an image is the contrast of an image. One way to improve the contrast is contrast stretching or histogram stretching. The concepts of image enhancements (i.e. increasing the contrast and features of an image) are important for recognition of objects and the important features of the generated synthetic images.

To match the generated synthetic images to the real camera images, the current dynamic range of grey values were mapped to the full range of grey values. That is the dynamic range of the generated images was compressed by linearly stretching the intensity values within the [0, 255] range. That is, each pixel a monochromatic image can have an 8-bit grey scale or intensity value ranging from 0 to 255. Figures 3.7, 3.9, and 3.11 show the histograms of the two images

(histograms of SCIs are shown in Figures 3.7a, 3.9a and 3.11a and for RCIs the histograms are shown in Figures 3.7b, 3.9b and 3.11b). It can be seen that few pixels have values near the extremes (no information on pixel location is given by the histogram) so the question is "should histogram equalization be applied to these images?". Histogram equalization (to obtain uniform histogram by redistributing the intensity distribution) can be used to increase the contrast across the whole dynamic range of the images. That is, histogram equalization is used for enhancing detail in low-range images and image enhancement is subjective and application specific (Gonzalez and Woods, 2002).

This thesis did not investigate how histogram equalization may affect the quality of the texture features extracted. However, in these histograms (i.e., in Figures 3.7, 3.9 and 3.11), it can be seen that the synthetic and real camera images display similar intensity distribution for the pixels. In other words, they are radiometrically similar for the three data sets. This also indicates that the generated synthetic camera images are equivalent to the real camera images. This similarity in the intensity distribution between the images facilitates the use correlation-based method which attempts to establish correspondence matching image intensities (discussed in Chapter 4).

3.3.4 Pre-processing of the RCI

The SCI is a monochromatic image and, thus, the RCI needs to be put into a compatible form. There are two options – the first is to convert the RGB RCI into

grey scale based on total intensity while the second is to use either the red, green or blue channels. An experiment using later elements of the fusion process was made to find out whether one of these was better than the other. The result of the experiment was inconclusive, so the grey scale images were adopted for the rest of the study.

3.4 Results and Analysis

As previously pointed out, the dynamic range values of the generated synthetic camera images were compressed to produce the final images shown in Figures 3.6, 3.8 and 3.10. Table 3.2 summarizes the details of the generated SCIs. In Figures 3.7, 3.9 and 3.11, histograms (showing the distribution of intensities in the grey scale image) of the generated images are presented. These histograms reveal that the pixel values are spread out between the minimum of 0 and maximum of 255.

3.4.1 The Synthetic Camera Image

Table 3.3: Details of the generated SCIs

Figure	Data Set	Resolution (Row x Column
3.5	**A**	2196 x 2009
3.6	**B**	1440 x 1851
3.7	**C**	2236 x 2086

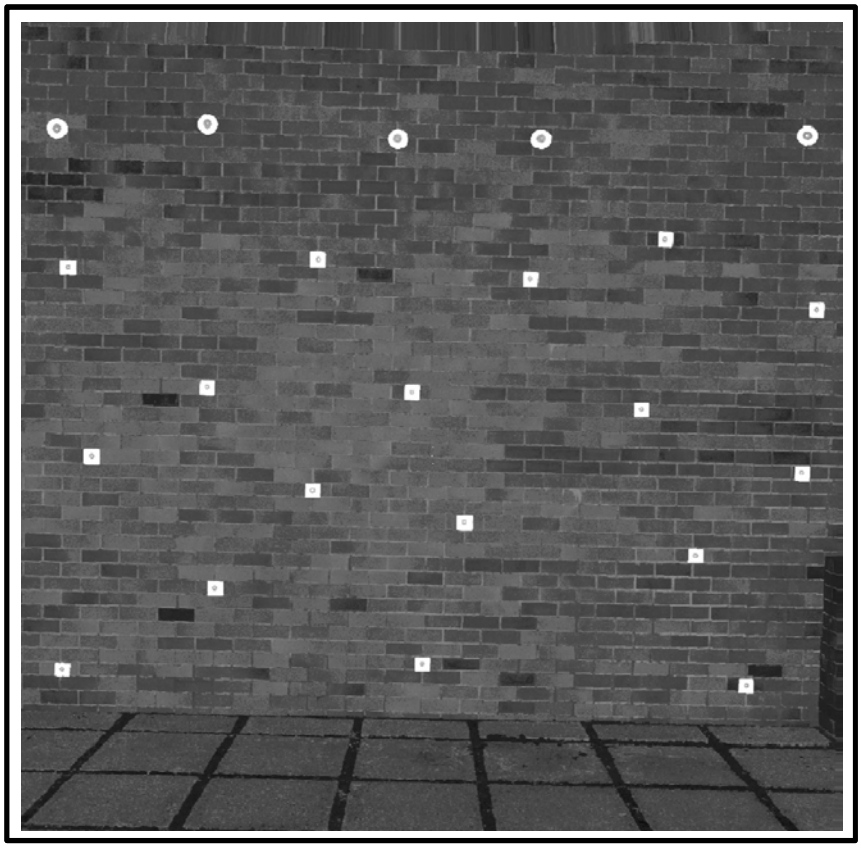


Figure 3.6: Generated Synthetic Camera Image of Data Set A

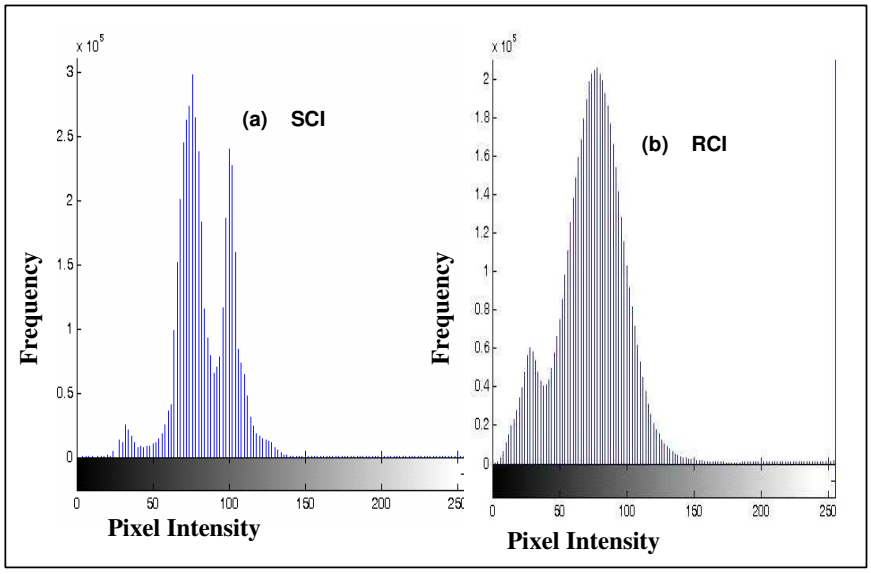


Figure 3.7: Histograms of Synthetic and Real Camera Images of Data Set A

Figure 3.8: Generated Synthetic Camera Image of Data Set B

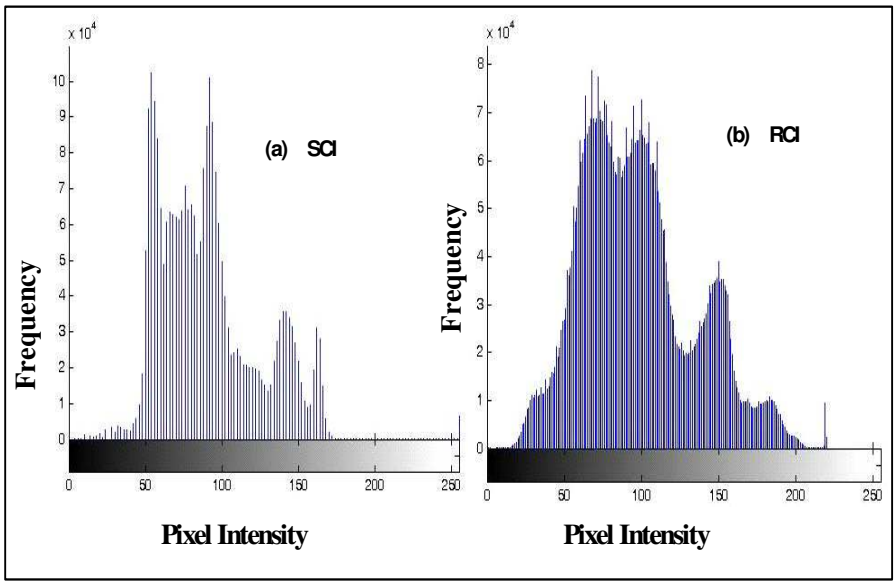


Figure 3.9: Histograms of Synthetic and Real Camera Images of Data Set B

Figure 3.10: Generated Synthetic Camera Image of Data Set C

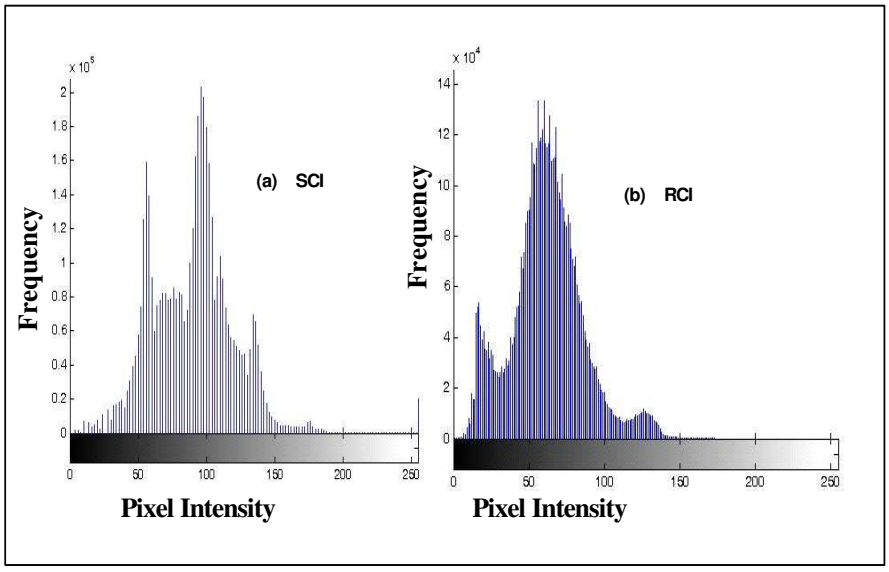


Figure 3.11: Histograms of Synthetic and Real Camera Images of Data Set C

3.4.2 Similarities and Differences between the Datasets

It is clear (as can be seen in Figure 3.12) that the geometric features in the SCI are easier to detect than those in the laser range data. This image offers a major advantage to interactively (controlled by a human operator) or automatically match conjugate points with the intensity images produced by a digital camera. However, there are differences in illumination, and reflectance as well as a lack of appropriate texture (Milian *et al*, 2002) between these images. Also, images from different sensors usually have their own inherent noise (Habib and Alruzouq, 2004). Furthermore, the automatic registration problem can be complicated, in this case, by differences in image resolution and scale, and low image quality (especially with the SCI).

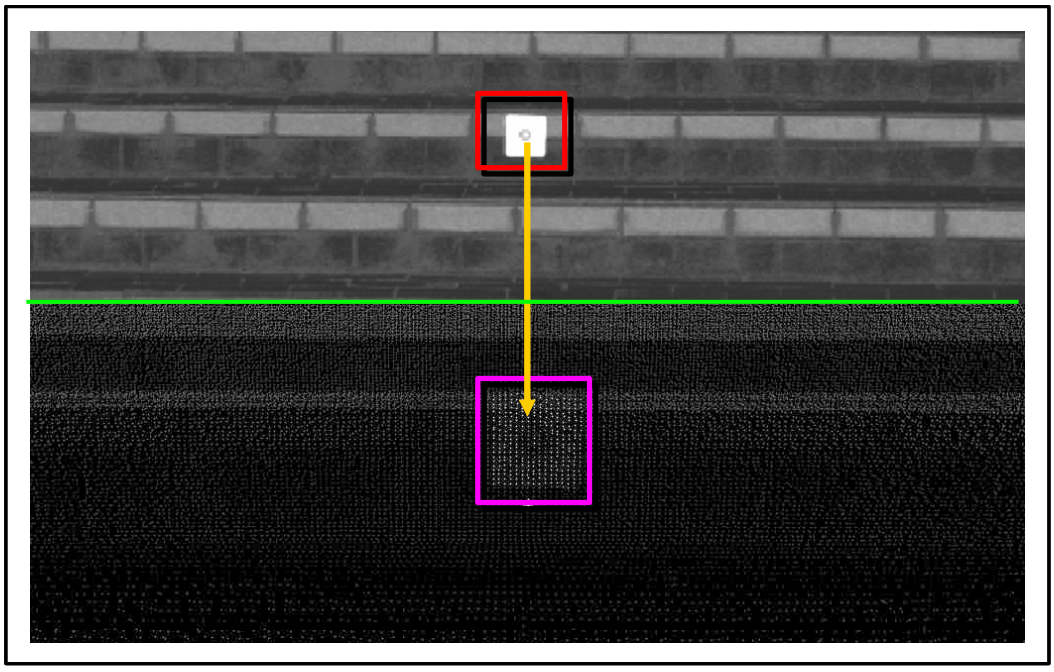


Figure 3.12: SCI (upper) and the Point Cloud (lower)

3.5 Summary

In this chapter, the strategy for terrestrial measurements (photogrammetry and laser scanning) has been discussed and the data sets used in this research has also been presented. In order to test both the fidelity of the SCI concepts and the correspondence algorithms, three different scenes were scanned with the laser scanner, and for each scan, a 2D photographic image was taken with a digital camera. The photogrammetric collinearity model was used to fuse data from the two sensors (digital camera and terrestrial laser scanner) to generate synthetic images. Synthetic camera images equivalent to that of a digital camera was created by interpolating the irregularly spaced synthetic data onto a grid at an even spacing and using the laser return pulse intensity.

These synthetic images of the three different data sets have been presented with their histograms showing the distribution of the pixels values. Additionally, the differences and similarities between the generated synthetic and real images have been discussed. Chapters 4 and 5 detail how the data set discussed in this chapter were used with developed algorithms to determine the geometric relationship between the digital camera and the laser scanner.

CHAPTER 4

AUTOMATIC HYBRID MATCHING FOR FEATURE LEVEL FUSION

4.1 Introduction

The fundamental task in this type of multisensor fusion – combining data from two distinctly different sensors (terrestrial laser scanners and cameras) is registering the two data sets together. At first glance, this may seem to be a trivial task as exemplified by the process of exterior orientation of a photograph via a set of control points to an object defined coordinate system. The quality of any products derived from such fusion will be directly influenced by the quality of the registration process. For this to be done, homologous points in the images need to be identified and then related to their object coordinates. This can be done manually or automatically. The actual results of both manual and automatic registration are given in the following chapter.

To perform automatic registration, three distinct steps are usually required: firstly to extract potential points, secondly to find those correspondences between the two images, and thirdly to determine the transformation parameters between them (this process is described in Chapter 5).

There are various strategies to solve this problem, the result for this research was a hybrid method based on the Harris corner detector to detect interest points, zero-mean normalized cross correlation for establishing the correspondence between points and RANSAC algorithm to filter out the false correspondences. The quality of the matching algorithm was evaluated.

4.2 Existing Matching Strategy for Feature Level Fusion

Many image processing problems involve the fundamental task of registration, which includes image matching of a pair of images. Usually, a matching procedure uses either feature-based methods (which use low-level features such as edges and corners) or area-based methods (which use pixel grey level values directly). Each of these were evaluated.

4.2 1 Area-Based Methods

The area-based or correlation-based method which attempts to establish correspondence matching image intensities usually adopts a window of points to determine a matched location (Bennamoun and Mamic, 2002, and Brown, 1992). Such methods include the sum of the absolute differences, the sum of squared differences, cross-correlation, and zero mean normalized cross correlation (Bennamoun and Mamic, 2002). Several correlation measures are reviewed, tested and compared in Giachetti (2000). All these measures use a sqaure window (i.e., a regularly sized pixel neighbourhood are compared to find the optimum match) of pixels as the basis for comparison. A regularly sized pixel neighbourhood is need

so that a point to be matched essentially becomes the centre of a small window of points, and this window is compared with similarly sized regions in the other image (Bennamoun and Mamic, 2002). In other words, area-based algorithms compare the intensity values within a regularly sized pixel neighbourhood centred at a point in one image with corresponding values in an identical regularly sized pixel neighbourhood centred at points in the other image.

The majority of the area-based methods have the limitation of registering only images with small misalignment, because they are sensitive to perspective distortion (Bennamoun and Mamic, 2002). If the orientation difference between the two images is large, the value of cross-correlation will be greatly influenced, and the correspondences between feature points, thus hard to derive. In order to solve the problem, it is necessary that images must be roughly aligned with each other initially to estimate the orientation parameters.

In this case, the correlation measures become unreliable when the images have multiple modalities and the grey-level characteristics vary (Dai and Khorram, 1999). This is because the measure of absolute difference in image intensity may not be meaningful as the pixel value represents the physical property that is unique to the imaging method. This thesis uses the zero mean normalized cross correlation (discussed in Section 5.5) as a similarity measure for the matching process

4.2 2 Feature-Based Methods

In contrast to area-based matching, feature-based matching uses symbolic descriptions of the image data for establishing correspondence. This symbolic representation consists of a set of geometric primitives (edges, contours, corners, line intersections). Feature-based techniques are more robust and are preferable in multisensor registration (Brown, 1992; Hsieh *et al*, 1997; Li *et al*, 1995). That is, feature-based matching is less variant against radiometric distortions. This is because feature-based matching does not work directly with image intensity values. This property makes feature-based methods suitable for situations when illumination changes are expected or multisensor analysis is demanded (Zitova´and Flusser, 2003).

Generally, feature-based methods are involved in feature selection and extraction and feature correspondence establishment. Image data contain a considerable amount of information, which forms a large search space for finding the information relevant to the domain of building objects. To reduce the dimensions of this search space, it is desirable to extract some basic features from the image.

4.3 Principles of Hybrid Matching for Feature Level Fusion

Area-based and feature based methods alone have some limitations and many high end photogrammetric systems now use a combination of the two techniques (Dias *et al*, 2002). Such an approach is known as hybrid matching. Such a methodology was developed for this research.

The subsequent section discusses the implementation of this hybrid approach. The initial results of this hybrid matching algorithm which consists of feature extraction process followed by the cross-correlation matching were published in Forkuo and King (2004a).

4.3.1 Detection and Extraction of Features for Matching

The automatic registration problem requires finding features (edges, corners) in one image and correlates them in another. For this thesis, Harris corner detector as proposed in Harris and Stephens (1988) was used to detect and extract corners in both images. This operator has been widely used and it has been shown to be robust to viewpoint changes (i.e. image rotations and translations) and illumination changes (Dufournaud *et al*, 2004; Rothfeder *et al*, 2003). However, the Harris corner detector is not invariant to changes in scale (Dufournaud *et al*, 2004). It uses a threshold on the number of corners extracted based on the image size. The number of corners detected in images is variable (Rothfeder *et al*, 2003). Once feature points are extracted from an image pair, correspondence matching can be performed.

4.3.1.1 Image Pyramids

One significant problem in this research was the difference in image size between the RCI and SCI. To account for different image sizes, the image pyramid was used. In this research, an intermediate step, an image pyramid, was computed for both images separately, before using any of the matching procedures.

An Image pyramid, which combines the advantages of both high and low resolutions of images, is a hierarchical structure composed of a number of levels (Figure 4.1) of the same image at different resolutions. This hierarchical approach involves the Gaussian smoothing and subsampling of the original images so that a "pyramid" of images at different resolutions is obtained (Bennamoun and Mamic, 2002).

The results of image pyramids are a series of reduced-resolution versions of the original images. Matching through the pyramid from coarse-to-fine reduces the search time for a match and increases the accuracy of the matching results (Bennamoun and Mamic, 2002; Heipke, 1997). Also, in the upper level images, features such as edges or regions of interest are more visible for correspondence analysis (Mikhail and Bethel, 2001). In other words, the lower levels of image pyramid provide detailed information.

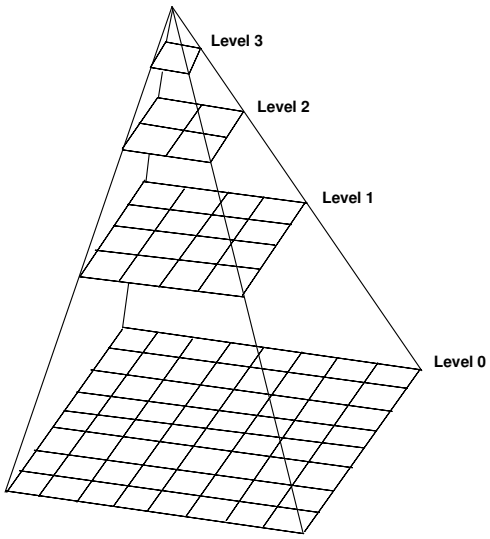


Figure 4.1: Image Pyramid

4.3.1.2 Extraction of Points of Interest

Image corner point features (also referred to as interest points in this thesis) are distinctive image points for which the signal change two-dimensionally (Schmid et al, 2000; Tissainayagam and Suter, 2004). In other words, interest points are the loci of two-dimensional intensity change (i.e. second-order features). Image corners impose more constraint on the motion parameters than do edges. Corners are also often more abundant than straight edges in real scenes or objects, making them ideal features to track in an indoor and outdoor environment.

A number of algorithms for interest points and corner detection have been reported in recent years (Schmid *et al*,2000; Tissainayagam and Suter, 2004). They can be divided into three categories: contour-based methods, parametric model based methods and intensity based methods (Schmid *et al*, 2000). Contour based methods involve extracting edges and then finding the points having maxima curvature, or searching for points where edge segments intersect. The Parametric model methods fit a parametric intensity model to the signal.

The intensity-based methods of algorithms search for point of interest or corners directly from the grey-level image. The Harris corner detector used in this thesis falls into the group of feature detectors based on intensity methods. Several interest points detectors are reviewed and compared in Schmid *et al* (2000). In the following section the Harris corner detector for the extraction of the aforementioned features is discussed.

4.3.1.3 Harris Corner Detector

Image corner detection is an important task in various computer vision and image understanding systems. Applications include motion tracking, object recognition and image matching. This algorithm is based on an underlying assumption that corners are associated with maxima of the local autocorrelation function (Tissainayagam and Suter, 2004).

The Harris corner detector has been widely used and it has been shown to be robust to viewpoint (i.e. image rotations and translations) and illumination changes (Dufournaud *et al*, 2004; Rothfeder *et al*, 2003). It should be noted that because this algorithm rely on spatial derivatives, image smoothing is often required to improve their performance. While improving the detection reliability, it has been shown that smoothing may result in poor localization accuracy (Tissainayagam and Suter, 2004). The Harris corner detector proceeds by forming an image gradient covariance matrix M (where M is a 2×2 matrix computed from image derivatives) corresponding an image I(x, y).

$$M = G(0,\sigma) \otimes \begin{bmatrix} \left(\frac{\partial I}{\partial x}\right)^2 & \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) \\ \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) & \left(\frac{\partial I}{\partial y}\right)^2 \end{bmatrix} \tag{4.3}$$

where:

$\frac{\partial I}{\partial x}$ and $\frac{\partial I}{\partial y}$ denote the image gradients in the x and y directions,

$G(0,\sigma)$ is a zero- mean Gaussian smoothing kernel of variance σ^2, and

$\otimes$ is the convolution operator.

The Harris corner detector then computes a 'cornerness' value R for each pixel in an image based on Rothfeder *et al* (2003)

$$R = \det(M) - k \times \text{trace}^2(M) \tag{4.4}$$

where:

det M denotes the determinant of the M (in Equation 4.3),

trace M is the sum of elements along the principal diagonal; and

k is a constant parameter that is usually taken to be 0.04 (Rockett, 2003; Harris and Stephens, 1998). R is the corner strength function at a point.

Maximum and minimum eigenvalues of the matrix M are then computed, and the maximum values indicate the corner position. That is, a corner is indicated when the two eigenvalues of the matrix M are large and similar in magnitude.

The Harris corner detector uses a threshold on the number of corner extracted based on the image size. The number of surrounding pixels required to calculate M is determined by the size of the Gaussian smoothing kernel. A typical size for neighbourhood is 5×5 pixels, though in practice it should be set as a percentage of

the image dimensions (i.e., the width and height of the Gaussian smoothing kernel should increase as the size of the source image increase), and its width should be odd, for reasons of symmetry.

The Harris corner detector was applied and tested over a large number of image pairs of data sets A, B and C. Figures 4.2, 4.3 and 4.4 are examples of corners detected with a Harris corner detector on image pairs (each image pair from each data set the remaining 3 image pairs for each data set are shown in appendix A3). It can be seen that the number of detected points are different for two the images of each data set and the differences in the number of interest points detected are due to the fact that the Harris corner detector is based on repeatability (which is defined by the image geometry) and information content (which is a measure of the distinctiveness of an interest point) of images.

The maximum number of corners that were detected by the Harris corner detector from the SCI was used a threshold to extract the same number of corners from the RCI. For example a maximum of 1000 corners were detected from SCI of data sets A and C (in Figures 4.2 and 4.4, respectively). This same number of corners (1000) was extracted from the RCI (as indicated in Tables 4.1 and 4.3). The same procedure was used for corner detection of data set B (in Figure 4.2), but here the maximum of corners was 992 (as indicated in Table 4.2). Once feature points were extracted from an image pair, correspondence matching could be performed.

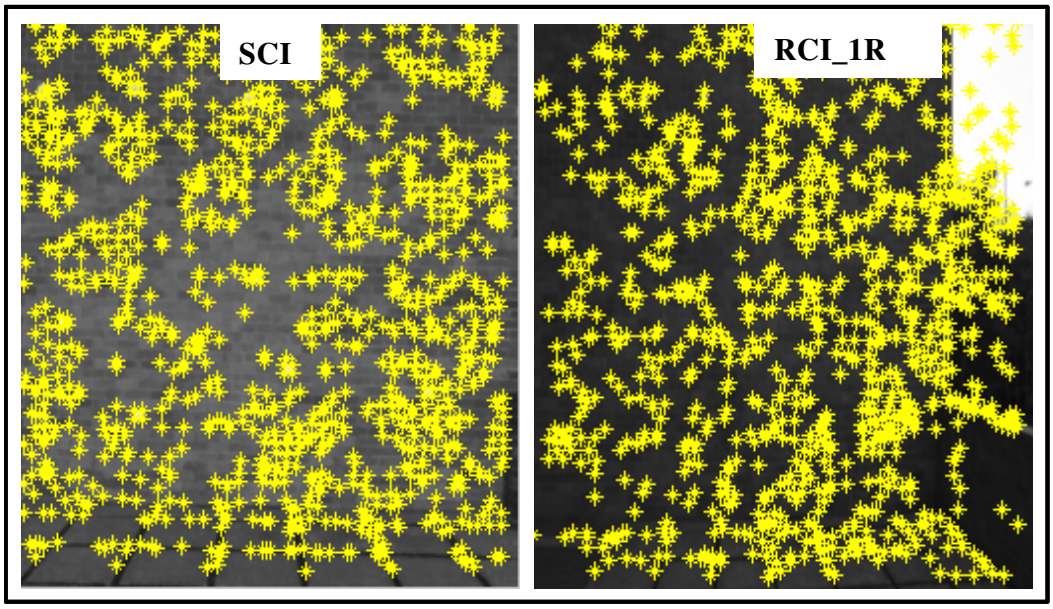


Figure 4.2: Example of Corner Detected with Harris Corner Detector on Model_1 of Data Set A.

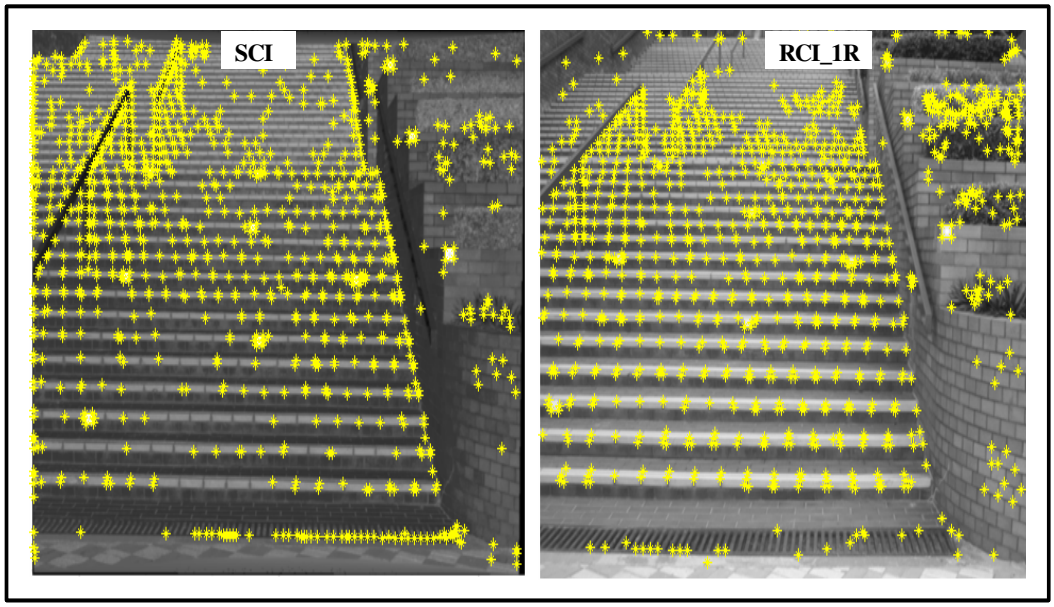


Figure 4.3: Example of Corner Detected with Harris Corner Detector on Model_1 of Data Set B.

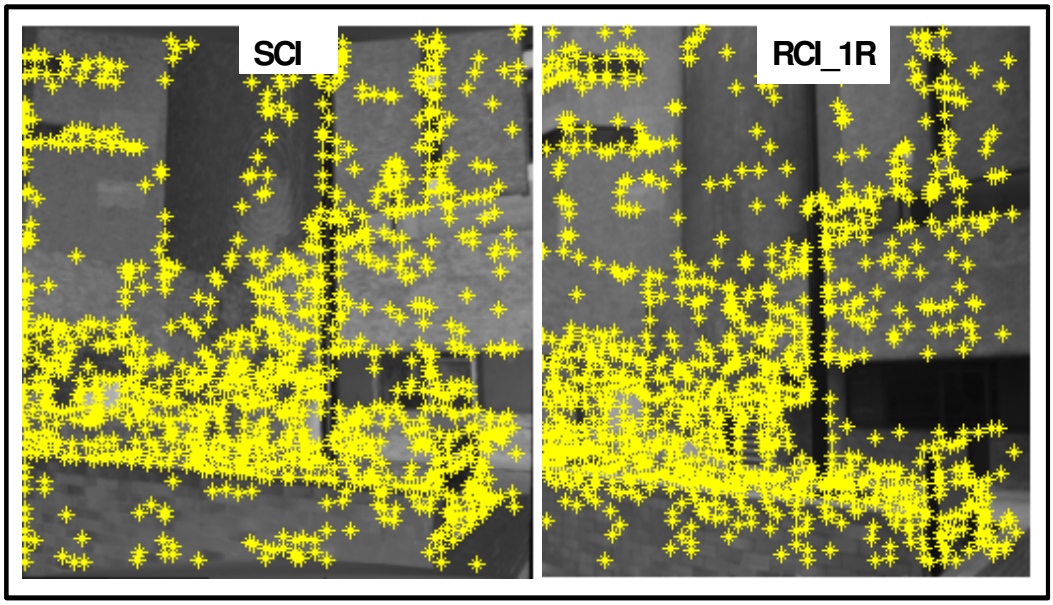


Figure 4.4: Example of Corner Detected with Harris Corner Detector on Model_1 of Data Set C.

4.3.2 The Correspondence Matching Algorithm

This section concentrates on determining the correspondence between the two sets of extracted interest points that were detected with a Harris corner operator. The most difficult part of the automatic matching is essentially the correspondence matching: given a point in one image, find the corresponding point in each of the other image(s). Although the automatic correspondence is not a problem for vertically oriented images, it is still a problem in the terrestrial case, and it is even most complex in a terrestrial multisensor case.

It can be observed that, since both image types are formed using similar mechanisms, the location of many objects in identifiable in each image. However, there are differences in illumination, perspective, reflectance as well as lack of appropriate texture (Milian *et al*, 2002) between these images.

Also, images from different sensors usually have their own inherent noise (Habib and Alruzouq, 2004). Furthermore, the automatic matching problem can be complicated by differences in image resolution and scale and low image quality (especially with the SCI).

Correspondence matching algorithm consists of two-step procedure. In the first step, matches are determined through normalized cross correlation. The initial matching can be used to eliminate points from both images that have no corresponding points. The second set up refines the initial corresponding points by using a robust estimator such as RANdom SAmple Consensus (RANSAC) algorithm (discussed in Section 4.5.1). Initial correspondence between these points is then established by correlating regions around the features. The similarity is then judged by the accumulated development of corresponding interest points in the two images (Rothfeder et al, 2003).

To match these features automatically, the zero mean normalized cross correlation (ZNCC) measure, which is invariant to varying lighting conditions (Lhaullier and Quan, 2000) is used. Mathematically, the ZNCC can be expressed as in Equation 4.7 (Crowley and Martin, 1995), where μ_S (Equation 4.5) is the mean of the template (of size M by N) and $\mu_R(i,j)$ (Equation 4.6) is the mean of the image neighbourhood (of size M by N). The match measure is intended to quantify the degree of 'similarity' between the sources. More precisely, the match value

(ZNCC(i, j)) (Equation 4.7) reflects the resemblance between the input images $S(i, j)$ and $R(i, j)$.

$$\mu_S = \frac{1}{MN} \sum_{m=0}^{M} \sum_{n=0}^{N} S(m, n) \tag{4.5}$$

$$\mu_R(i, j) = \frac{1}{MN} \sum_{m=0}^{M} \sum_{n=0}^{N} R(m + i, n + j) \tag{4.6}$$

$$ZNCC(i, j) = \frac{\sum_{m=0}^{M} \sum_{n=0}^{N} (S(i, j) - \mu_S)(R(i + m, j + n) - \mu_R(i, j))}{\sqrt{\sum_{m=0}^{M} \sum_{n=0}^{N} (R(i + m, j + n) - \mu_R(i, j))^2 \cdot \sum_{m=0}^{M} \sum_{n=0}^{N} (S(i, j) - \mu_S)^2}} \tag{4.7}$$

This method uses a small window around each point to be matched (this point becomes the center of a small window of grey level intensities), and this window (template) is compared with similarly sized regions (neighbourhood) in the other image (Rothfeder *et al*, 2003). In other words, the ZNCC method is based on the analysis of the grey level pattern around the point of interest, and on the search for the most similar pattern in the successive image (Giachetti, 2000). Each comparison yields a score, a measure of similarity. The match is assigned to the corner with highest matching score (Smith *et al*, 1998).

Selection of an appropriate parameters such as correlation window size is a central problem with the ZNCC techniques. An experiment was performed with various parameters to find out suitable ones.

The results showed that by selecting a patch size (correlation window of 17 x 17 pixels), a neighbourhood window size of 50 x 50 pixels and correlation coefficient threshold of 0.8, suitable matching results were obtained. That is, the matching process reduces the number of detection of false correspondence pairs. However, the number of mismatches (referred to as outliers) may be quite large.

This occurs in particular when some corners cannot be matched. Also, there are likely to be several candidate matches for some corners which are very similar (Smith *et al*, 1998). In Figures 4.5, 4.6 and 4.7, matches are shown by the line linking matched points to their position in the other image (clear mismatches can be seen). The summary of the matching results for data set A, data set B and data set C are shown in Tables 4.1, 4.2 and 4.3, respectively.

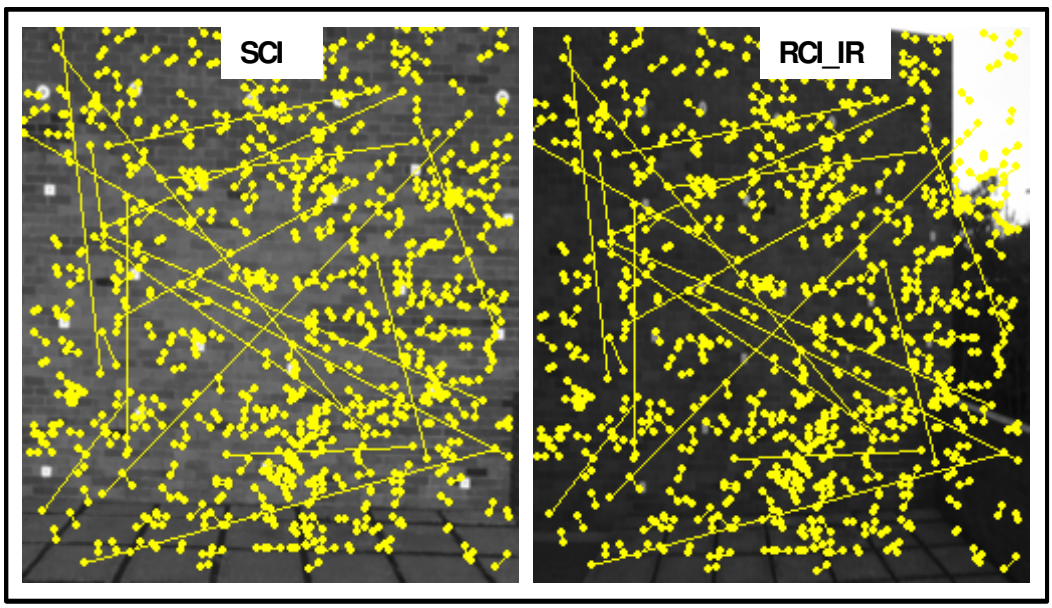


Figure 4.5: Example of Detected Correspondences of Model_1 of Data Set A with ZNCC.

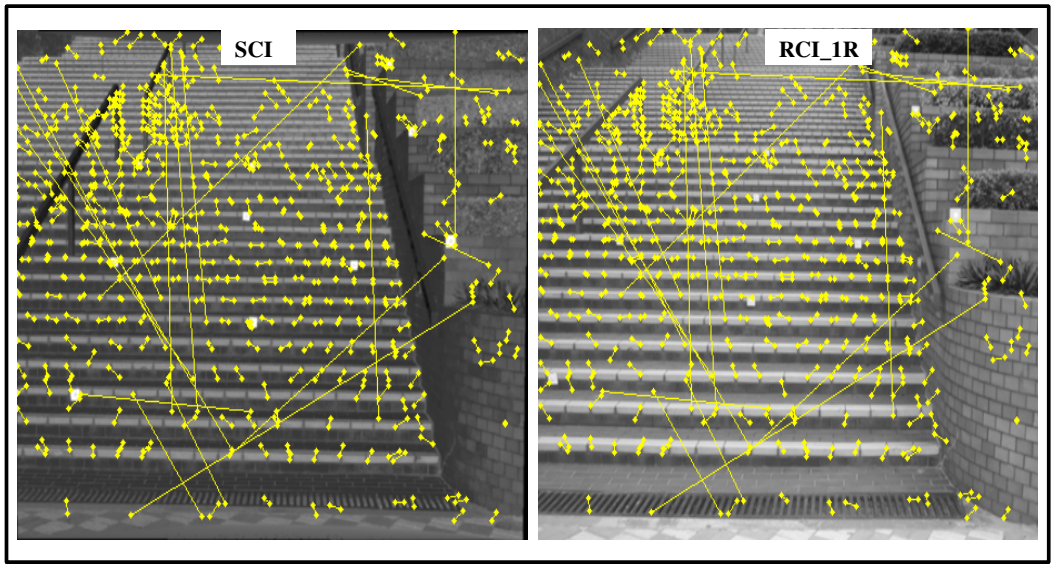


Figure 4.6: Example of Detected Correspondences of Model_1 of Data Set B with ZNCC.

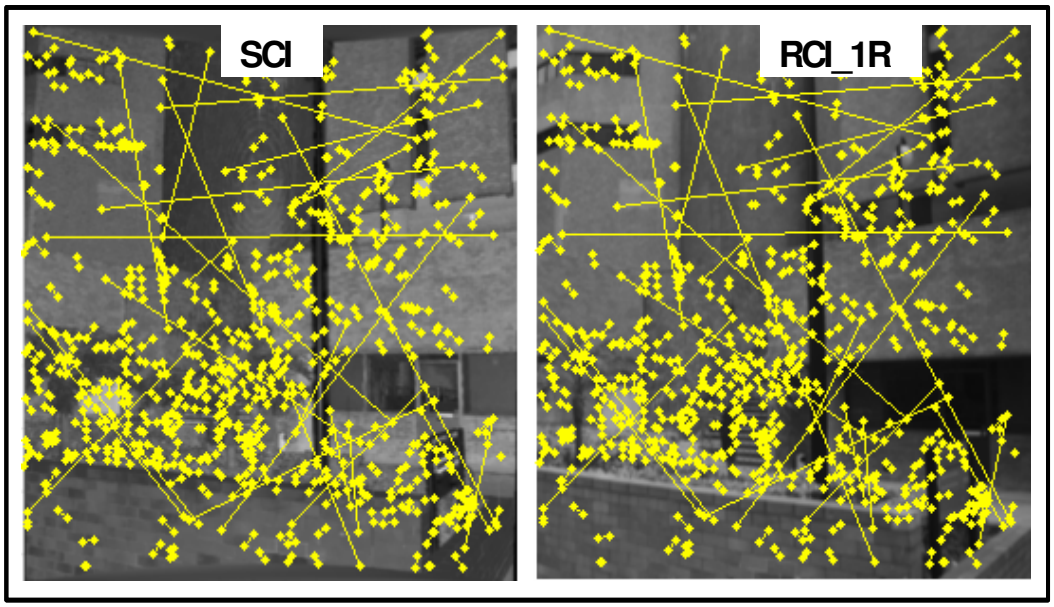


Figure 4.7: Example of Detected Correspondences of Model_1 of Data Set C with ZNCC.

4.3.2.1 Robust Matching for Filtering False Correspondences

The most commonly used robust estimators include M-estimators, Least-median-squares (LMedS), Maximum A Posterior Sample Consensus (MAPSAC), and RANSAC ((Dufournaud *et al*, 2004, Armangue and Salvi, 2003). In contrast to most of robust algorithms which attempt to maximize the quantity of data used to identify a solution, RANSAC algorithm of Fischler and Bolles (1981) robustly searches for suitable solutions directly using the data, repeatedly constructing solutions from randomly sampled minimum subsets of data which are not related to any concept of an error surface (Lacey *et al*, 2000). This thesis, therefore, uses the RANSAC algorithm to refine the results of the cross correlation (discussed in 4.5). The implementation details are discussed in Capel and Zisserman (2003); Fischler and Bolles (1981), and Hartley and Zisserman (2000).

The RANSAC technique is based on randomly selecting the set of points used to compute an approximation of the fundamental matrix (F) by a linear method (Armangue and Salvi, 2003). Thus, RANSAC robustly determines the F from the correspondence matches (discussed in Section 4.5), without requiring the intrinsic and extrinsic camera parameters (Hartley and Zisserman, 2000).

If x is a point in one image and x' a point in another image, then the image points satisfy the relation $x'^{T}Fx = 0$ (details of the computation are described in Hartley and Zisserman (2000). Once the F matrix is determined, one can reconstruct the epipolar geometry which is the intrinsic projective geometry between two views

(Hartley and Zisserman, 2000, Pollefeys *et al*, 2000) and matches can be found by searching along epipolar lines. In most circumstances, the estimated F should be a rank 2 matrix in order to model the epipolar geometry with all the epipolar lines intersecting in a unique epipole (Armangue and Salvi, 2003).

Computing the epipolar geometry from the set of potential matches through least squares generally does not give satisfying results (Pollefeys *et al*, 2000). The obvious advantage of using epipolar geometry is that the conventional 2D correlation can be completely substituted by near 1D correlation along the epipolar lines. The matches (i.e., the results of the correspondence matching in Section 4.5) are refined using with RANSAC algorithm. This algorithm allows the user to define in advance the number of potential outliers through the selection of a threshold.

A non-linear minimization is used to fit F to a large number of points. The best solution is that which maximizes the number of points whose residuals are below a given threshold. Once outliers are removed, the set of points identified as inliers may be combined to give the final solution (RANSAC inliers). The results are shown in Figures 4.8, 4.9, and 4.10. These inlying correspondences are used in the exterior orientation process. The summary of matching results is shown in Tables 4.1, 4.2 and 4.3 for data sets A, B and C respectively. These tables show, for each model, the number of features detected (i.e., number of potential matches), the

number of correspondences (matches) obtained, the final number of matches, and the percentage of inliers.

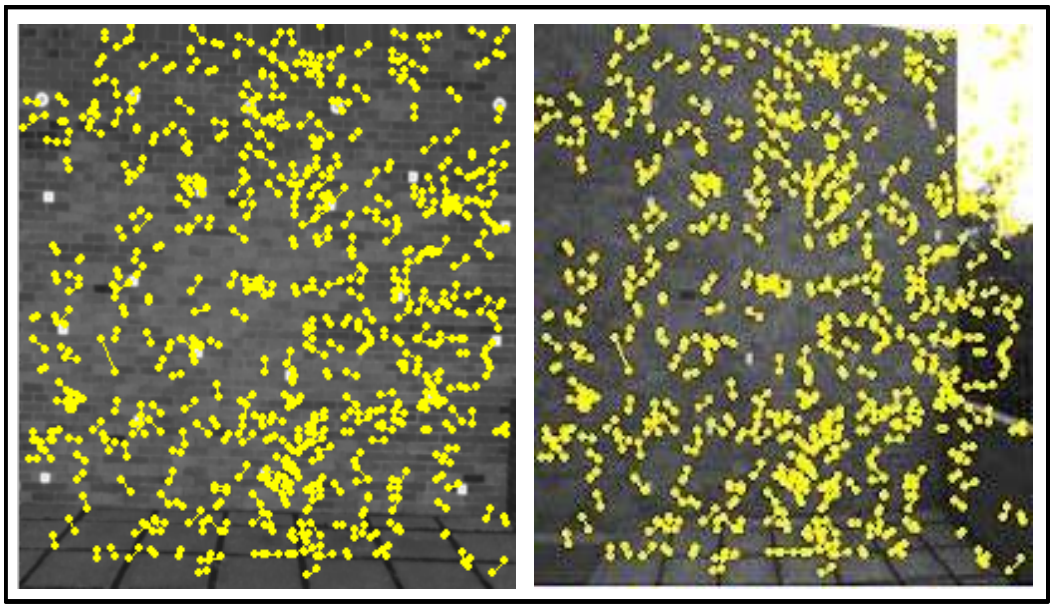


Figure 4.8: The Final Verified Detected Correspondences (RANSAC inliers) of Model_1 of Data Set A

Table 4.1: Summary of Matching Results of Data Set A

Data Set A	Model_1	Model_2	Model_3	Model_4
No. of features detected	1000	1000	1000	1000
Mutual correspondence Obtained	474	434	449	414
RANSAC inliers (valid matches)	317	297	265	280
% of valid matches	67	68	59	68

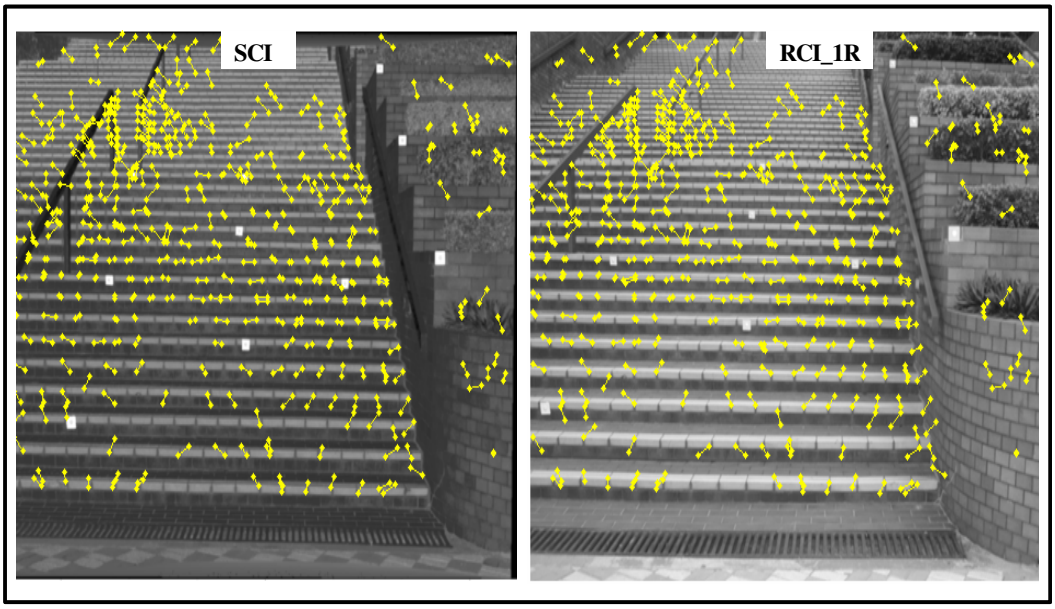


Figure 4.9: The Final Verified Detected Correspondences (RANSAC inliers) of Model_1 of Data Set B

Table 4.2: Summary of Matching Results of Data Set B

Data Set B	Model_1	Model_2	Model_3	Model_4
No. of features detected	992	992	992	992
Mutual correspondence Obtained	383	366	383	363
RANSAC inliers (valid matches)	248	256	276	267
% of valid matches	65	70	72	74

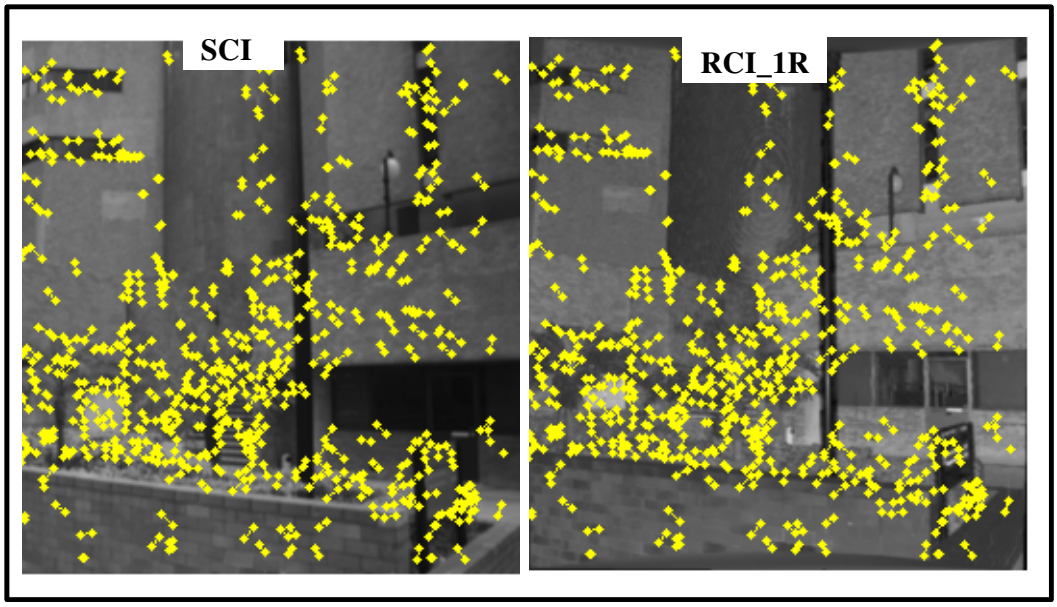


Figure 4.10: The Final Verified Detected Correspondences (RANSAC inliers) of Model_1 of Data Set C

Table 4.3: Summary of Matching Results of Data Set C

Data Set C	Model_1	Model_2	Model_3	Model_4
No. of features detected	1000	1000	1000	1000
Mutual correspondence Obtained	385	402	420	393
RANSAC inliers (valid matches)	256	296	279	241
% of valid matches	66	73	66	61

4.4 Evaluation of the Methodology

The proposed matching algorithm was tested on three different types of data sets of different scenes. The evaluation of the matching methodology is based on assessing the quality of the detected interest points and the feature matching algorithms.

4.4.1 Assessing the Quality of the Detected Interest Points

Having found corners in one image, the same corners should be found and matched in successive images. The ability to consistently find and match corners in this way relies on the corners being accurately localized (Noble, 1998). That is, the calculated image-plane position of a corner, given by the detector, should be as close to the actual position of the corner as possible. The Harris corner detector used in this thesis is largely controlled by setting parameters for 'cornerness' and Gaussian sigma (which specifies how much to smooth the image before processing) and the relative minimum threshold (which specifies the minimum corner strength).

While improving the detection reliability, it has been shown that smoothing may result in poor localization accuracy (Tissainayagam and Suter, 2004). The inherent disadvantage of this method is that the threshold is difficult to choose in practice and often arbitrarily assigned (Bennamoun and Mamic, 2002). These techniques depend largely on the threshold to determine whether or not a match is valid.

4.4.2 Assessing the Quality of the Matching Algorithm

This hybrid algorithm implemented is sensitive to: number of features; neighbourhood size; and window block size (for ZNCC). The size of the matching neigbourhood has a significant impact on the quality of the matches. However, there is a number of techniques which can be used to recognize incorrect or

invalid matches. Most of these are discussed in Bennamoun and Mamic (2002). In this research, the validation techniques used for the area-based matching are: the value of the matching score, and the left-right consistency checking. With the value of the correlation score, the score may be compared with respect to an absolute threshold.

A match whose score is too low (in the case of the ZNCC) was considered unreliable. With the correlation method, the assumption was that the grey level pattern is approximately consistent between the successive images (with no perspective effects) and that the local texture contains sufficient unambiguous information (Giachetti, 2000).

The left-right consistency checking was also implemented to remove the invalid matches caused by occlusion. That is, this test was used to identify invalid matches which may result from an occlusion. This technique involves reversing the roles of the two images and performing matching a second time. If the two sets of matches are the same, the match is considered consistent; otherwise it is flagged as inconsistent. Figures 4.6, 4.7, and 4.8 show that the value of the correlation score and the left-right consistency checking could remove a significant number of invalid matches. However, as can be seen, there was a number of invalid matches.

Also, selection of an appropriate correlation window size is a central problem with the ZNCC techniques so that the question is "what is the ideal size of the windows used to perform the matching?". The correlation window must be large enough to include sufficient information for matching, but small enough to avoid the effects of projective distortions (Bennamoun and Mamic, 2002). The size of this window also affects the time it takes to compute the matches. The window size selected was 17 x 17 pixels. The RANSAC algorithm was used to robustly filter out the invalid matches and it seemed to produce acceptable results (as can be seen in Figures 4.9, 4.10 and 4.11)

4.5 Summary

The chapter described the automation of the image matching process. Area and feature based method for matching were evaluated. A hybrid method that integrates the merits of both area and feature based approaches was adopted. An image pyramid was computed for both images separately, before using any of the matching procedures. A Harris corner detector with appropriate threshold value was first used to extract feature points in both images. The extracted features points were then matched with zero-mean normalized cross-correlation similarity metrics to find correspondences between the SCI and RCI.

The algorithm computes similarity scores for every pixel in the image by taking a fixed window (i.e., the algorithm compares the intensity values within a square window centred at a point on one image with the corresponding value in an

identical square window centred at points on the other image). A validity check was implemented by doing the correlation twice by reversing the roles of the two images. The RANSAC algorithm was used to filter out false (or incorrect) correspondences.

CHAPTER 5

MULTISENSOR IMAGE REGISTRATION

5.1 Introduction

Image registration can generally be defined as a mapping between two or more (images acquired at different times, different sensors or different viewpoints) both spatially and with respect to intensity (Brown, 1992; Rahman *et al*, 200). Mathematically, the problem of registration can be expressed as (Brown, 1992):

$$I_2(x', y') = g(I_1(T(x, y))) \tag{5.1}$$

where :

I_1 and I_2 are 2-dimensional images (indexed by x and y),

T is a transformation which maps the two spatial coordinates,

x and y to two new spatial coordinates x' and y', and

g is a one-dimensional intensity or radiometric transformation.

The result of the registration is a mathematical description of how the multiple images overlap, and is a fundamental concept used in photogrammetry.

Its usefulness can be seen in many applications including 3D mapping of the land and sea surface, identifying and mapping different types of land use, measuring and analyzing natural and agricultural vegetation, matching stereo images to recover shape for autonomous navigation, aligning images from different medical modalities for diagnosis, developing coded targets for automated identification, and coordinates determination of marked points for 3D model reconstruction.

As previously mentioned, to perform registration, three distinct steps are usually required: extraction of potential points, correspondence matching between the images, and determination of the transformation parameters. In the previous chapter, features were extracted from the generated synthetic camera images, and the photographic images and the correspondences (candidate registration points) between them were established (i.e., matches between the SCI and RCI were obtained) for use in the registration process. Finding the parameters of the optimal spatial or geometric transformation is generally a key problem to registration.

This chapter now describes the process of determining the parameters of the geometric transformation (to calculate the relative rotation and translation between the two sensors) of the SCI and RCI relative to the point cloud. The implementations of both manual and automatic methods for the orientation computation (exterior orientation parameters) are described in this chapter. The manual method uses the existing photogrammetric application software for the computation of the transformation parameters.

In this research, for the automatic computation of the transformation parameters (or exterior orientation parameters) two methods were investigated. Each of these methods is discussed in this Chapter. The first method which, works in the frequency domain, uses the Fast Fourier Transform (FFT) approach and the second method uses the photogrammetric bundle adjustment. The latter approach uses candidate registration points (discussed in Chapter 4) to determine the relative rotation and translation between the two sensors (laser scanner and digital camera).

The bottleneck of automation of image orientation is the identification of control points for orientation (Labe *et al*, 1996). The solution which has been developed and implemented (discussed in Section 5.3) for the automation of the exterior orientation is the use of Point Cloud Visual Index (PVCI). The automatic exterior orientation without the need for control point is a possibility as well the terrestrial applications discussed in Chapter 7. The assessment of the performance of each method was based on the accuracy of the computed object point coordinates as it is this criterion which is generally of greatest interest in close-range analytical photogrammetry.

5.2 Manual Exterior Orientation with StereoMaker Software

As previously pointed out, the images used in this study are images of the same scene acquired by different sensors and at different view points or locations. These multisensor and multi-view images of the same scene may have relative

translation, rotation, scale, and other geometric transformations between them, and the purpose of this is to manually establish the correspondence between them and to determine the geometric transformation or the orientation parameter that aligns one image with the other.

For this manual measurement, it is necessary to understand key issues such as geometric quality, and both spatial and geometric resolutions of the generated synthetic camera images. In addition, since there was no known value upon which to test the exterior orientation results against as there was no calibration data for either camera system, the results of the manual measurement (i.e., the manual orientation parameters) are used to compare and evaluate the proposed automatic orientation strategy for the fusion of the multisensor data (discussed in Chapter 6). The results show that the manual errors were generally larger than those of the automatic.

5.2.1 Implementation of the Manual Exterior Orientation Process

As explained earlier, the prime objective of this research was to automatically align 2D image and 3D object. This section presents the implementation of the orientations processes with the 3DMapper software StereoMaker using the Scanner Cartesian Coordinate (SCC) system as the object space coordinates. StereoMaker uses a bundle adjustment to process the image measurements to produce the object space point coordinates of all the measured points.

Two of the three data sets discussed in Chapter 3 (data sets A and B) were used in this manual orientation process. Four models or image pairs (i.e. model_1 (SCI, RCI_1R; model_2 (SCI, RCI_1L); model_3 (SCI, RCI_2L and Model_4 (SCI, RCI_2R) of each data set were used to test how well the generated SCI can be related to the real camera images.

The StereoMaker software offers the ability to perform affine, projective and conformal transformations for interior orientation and displays the root mean square errors of the residuals in pixels. Examination of the various transformation results showed that conformal transformation (i.e., transformation that preserves both angles and shapes) was the most suitable. Using the scanner point cloud to define a coordinate system, the exterior orientation was performed by a bundle adjustment.

Twenty-two control points (fine scan Cyra target coordinates) were distributed through the scan for data set A, and ten for data set B. These Cyra targets were used as control points for the orientation, their coordinates were obtained from identifying the center of each target in the point cloud. The performance of this manual method is based on the assessment of the root mean square errors of the object point coordinate. The exterior orientation parameters are presented and discussed in Chapter 6.

Initial implementation and results of manual multisensor image registration, which includes, interior, relative and absolute orientations using two different types of software for comparison purposes, have been discussed in Forkuo and King (2003).

5.2.2 Evaluation of the Manual Exterior Orientation Process

There are several factors that influence the accuracy of the orientation process. The quality of this photogrammetric measurement depends mostly on scale, number of images, image resolution, the quality of camera description the accuracy and precision of the image point measurements and the accuracy of the control points (number and distribution of fine scan Cyra target coordinates).

In this thesis, the effect of camera calibration on the overall orientation is not investigated. That is, the RCI was not calibrated (assumed principal distance, principal point location and no lens distortion correction). Assessment of the performance of the manual orientation is based on the accuracy of the computed object coordinates as it is this criterion which is generally of interest in close-range analytical photogrammetry.

5.2.3 Accuracy of Manual Measurements

Accuracy measures used to verify the geometric accuracy quality are the individual components of control point residuals and the root mean square (RMS) errors.

These residuals represent the differences between the object point coordinates as produced by each StereoMaker orientation, and the known object point coordinates from the laser scanner data (laser scanner coordinates).

5.2.3.1 Accuracy of Data set A

A graph of the residuals for individual control points in each model (image pair) for data set A is given in Figure 5.1. It can be seen from this figure that the residuals vary from 0.000-0.009mm for all models in X, Y, and Z directions. The maximum residual for all the models for data set A is 9mm which was significantly smaller than the nominal positional errors in the XY plane ($2Se_{xy}$ = ±12mm) for a single point of the laser scanner. This maximum residual of 9mm for the Z is slightly greater than the nominal range accuracy ($2Se_{xy}$ = ±8 mm) for a single point.

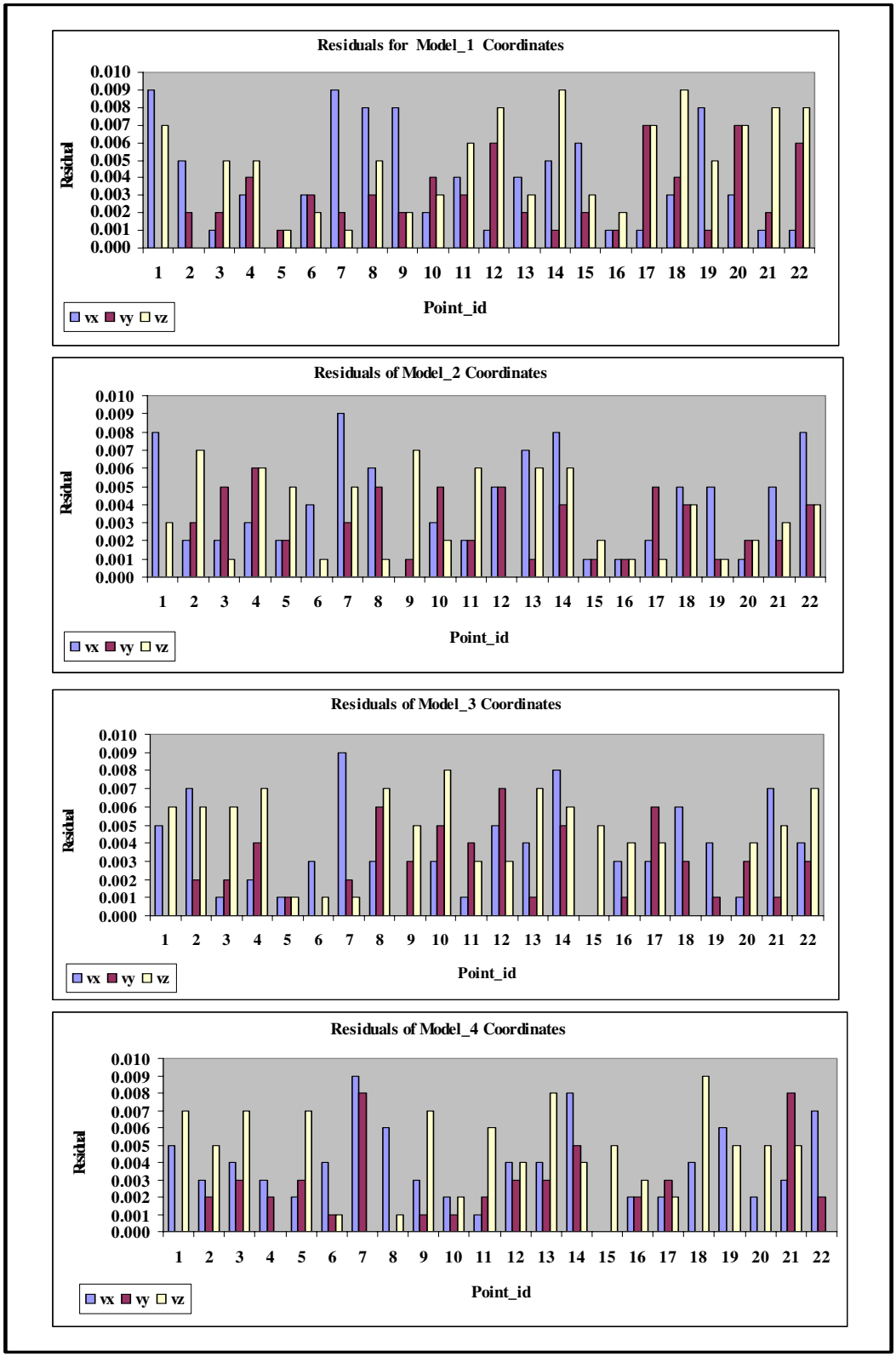
What is most evident in the results is that the control point 7 (point_id 7) has the largest residual (the software produces the computed coordinate with comparatively poorer accuracy) for all the models in data set A. The remaining residuals are small indicating the accuracy measures of the manual bundle adjustment showed good correspondence (with maximum residual of 8mm for X, Y and Z) to the differences between the determined and the predefined object point coordinates.

Table 5.1 summarizes the results from StereoMaker software with overall Root Means Square (RMS) errors of the measurements for all the models in X, Y, and Z directions. It can be seen that the RMS errors (varying from 0.003-0.006m) for all the four models is less than one pixel (pixel size of 8 micrometers).

The maximum RMS error in X and Y is even significantly smaller than the nominal positional error of the scanner ($2Se_{xy} = \pm 6$mm) and the RMS error in Z is also smaller than $2Se_z$ (the nominal range accuracy of ±8 mm).

Table 5.1: RMS Errors of Computed Object Coordinates for Data Set A

RMS errors(m)	Model_1	Model_2	Model_3	Model_4
r.m.s_x	0.005	0.005	0.004	0.004
r.m.s_y	0.003	0.003	0.003	0.003
r.m.s_z	0.006	0.004	0.005	0.005



Figures 5.1: Residuals (in metres) of Computed Object Coordinates for Data Set A

5.2.3.2 Accuracy of Data Set B

It can also be seen from Figure 5.2 (residual errors for data set B) that the values for the control point residual range between 0.000 and 0.009m for all the four models. The maximum RMS errors in X and Y were less than ±12 mm ($2Se_{xy}$). The largest error was 0.009m and it was contributed by point 6 (with point_id of 6) of model 2. The remaining X and Y residual for all the control points appears to be generally lower than 9mm for all the models. As can be seen, the maximum z-coordinate residuals for all models for were generally higher with the maximum value being greater than ±8 mm ($2Se_z$). The z-coordinates of point_id 1 (model_2), point_id 3 (model_3), point_id 4 (of model_2 and model_3), point_id 7 (model_1) had the maximum residuals greater than ±8 mm ($2Se_z$) and further analysis is need to explain these high residual values.

Also, the overall RMS errors for all the four models are provided in Table 5.2. What is most evident in the results is that, most of the models have a RMS error of approximately 2-4 mm (in X and Y and 5-8 mm (in Z). The RMS errors in X and Y were even less the ±6 mm (Se_{xy} = ±6 mm) and the maximum RMS error of the Z was equal to the ±8 mm ($2Se_z$). In each model for this data set, the RMS Z coordinates error was larger than the RMS X and Y coordinate errors. It should be noted that, in general for data sets A and B the Y coordinate for all the models showed the smallest RMS error. This trend did not occur in same data sets (data sets A and B) for the automatic measurements and it therefore, requires further investigations for a more conclusive evaluation.

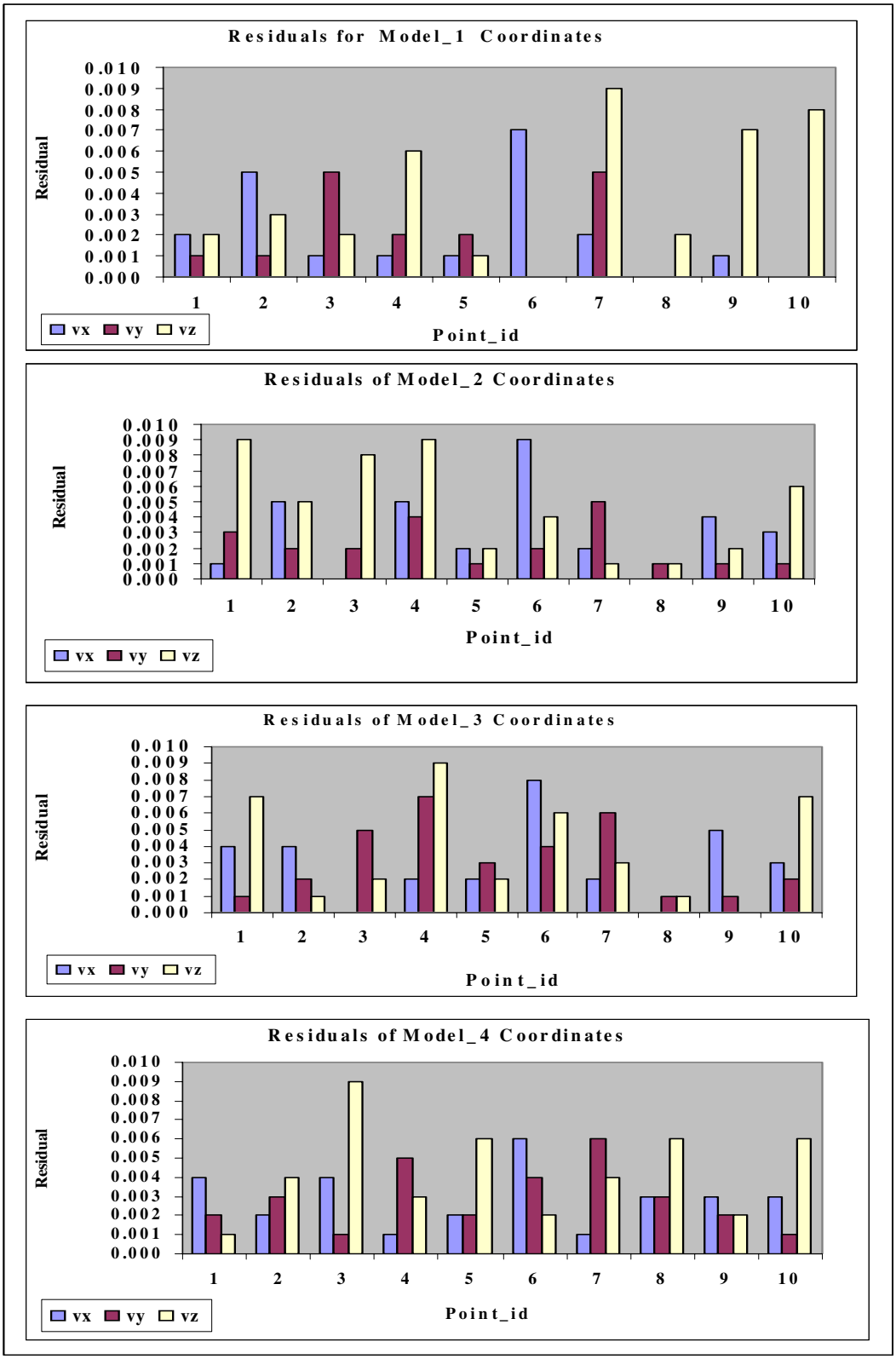


Figure 5.2: Residuals (in metres) of Computed Object Coordinates or Data Set B

Table 5.2: RMS Errors of Computed Object Coordinates for Data Set B

Root Mean Square Errors(m)				
	Model_1	**Model_2**	**Model_3**	**Model_4**
r.m.s_x	0.003	0.004	0.004	0.003
r.m.s_y	0.002	0.002	0.004	0.003
r.m.s_z	0.005	0.006	0.005	0.008

5.3 Automatic Exterior Orientation

Several algorithms for automatic registration have been proposed and successfully tested. In this research, two algorithms have been developed and implemented. The first is the frequency-based algorithm (the Fast Fourier Transform) which uses the frequency domain information to determine the exterior orientation parameters. The second is the photogrammetric bundle adjustment which uses both feature and area based algorithms to determine correspondence between two images and then computes the exterior orientation parameters. The initial results of the automatic exterior orientation process were discussed in Forkuo and King (2004b).

5.3 1 Evaluation of Frequency-Based Methods

The Fourier domain approach can be used to match images that are translated, rotated and scaled with respect to one another by making use of the phase shift property of Fourier transforms (Reddy and Chatterji, 1996).

If an acceleration of the computational speed is needed, or if the images were acquired under varying conditions or they are corrupted by frequency-dependent noise, then Fourier methods are preferred rather than the correlation-like methods (Zitova and Flusser, 2003). Fourier methods differ from other registration methods because they search for an optimal match according to the information in the frequency domain.

The phase correlation method (i.e phase correlation method that uses Equation A2.3 to extract the phases of the cross-power spectrum of two images) is based on the Fourier Shift Theorem (Zitova and Flusser, 2003) and was originally proposed for the registration of translated images. It computes the cross-power spectrum of the sensed and reference images and looks for the location of a peak in its inverse. This method shows strong robustness against the correlated and frequency dependent noise.

FFT-based approaches to image registration have been considered for many years. For example, Kuglin and Hines (1975) developed a method called phase correction by using certain properties of the Fourier transform; De Castro and Morandi (1987) introduced an extension of the phase correlation to rotation as well as shift; Reddy and Chatterji (1996) improved on De Castro's algorithm by greatly reducing the number of transforms that need to be performed. Applications of the extended algorithm in remote sensing (SPOT images) and medical imaging are described in Chen *et al* (1994).

The registration method based on FFT relies on the Fourier shift theorem, which states that a shift in the coordinate frames of two functions or images is transformed in the Fourier domain as linear phase differences. That is, the shift theorem guarantees that the phase of the cross-power spectrum is equivalent to the phase difference between the images (Reddy and Chatterji, 1996). The necessary theory for pairwise image registration using Fourier translation, rotation and scaling properties is summarized in Appendix A2. In this research, the FFT-based algorithm for automatic pairwise image registration was implemented using the Matlab programming language.

The matching parameters (rotation, scale and translation) for the test multisensor images are summarized in Appendix A2 .The "P-value" indicates the peak value, whose location gives the amount of rotation and the scale in the log-polar plane. Using the FFT theory, matches are considered valid only if the peak value of the inverse FFT of the phase difference is greater than 0.03 (Reddy and Chatterji, 1996). Theoretically, for exact matches this value should be equal to 1.0. However, the presence of dissimilar parts and noise in the test images reduces the peak values (as seen in Tables A2.1, A2.2 and A2.3). It has been observed from several experiments that if the peak value is less than 0.03, then the match will become unreliable.

Also, the algorithm requires images of the same type (from the same sensor) that were acquired in almost the same season (Xie *et al*, 2000). It works well with

vertically oriented images. As pointed out in Section 2.4.1.1, the sensors used in this research provide independent datasets, which differ in resolution, field of view (FOV), scale, illumination and reflectance. The images are also convergent and this explains why the algorithm (i.e. the conventional phase correlation method) did not work well in multisensor images such as the ones used in this thesis.

5.3.2 A Point Cloud Visual Index for Automatic Exterior Orientation Process

In this context, Point Cloud Visual Index-PCVI establishes a link between the pixel intensity values of the SCI and the corresponding sampled 3D point data (which can be referred to as the object coordinates). The two terms *visual* and *index* in the PCVI signify the intensity information (visual) of the point cloud and the mapping of the pixel of the SCI to its object coordinate (indexing) of the point cloud, respectively.

5.3.2.1 The Concepts of Point Cloud Visual Index (PCVI)

The core of the PVCI is the generated SCI- a visual, 2D representation of the TLS point cloud. With the SCI, there is no need to identify control point as each point in the SCI can be used as a control point. In the presence of point cloud data with return signal intensity information, large blocks of RCIs can be automatically oriented. For this to be done homologous points in the images need to be identified and then related to their object coordinates.

By linking the object coordinates of the point cloud to the pixels of the generated SCI, a database consisting of image coordinates and object coordinates of the point clouds is obtained.

This procedure is similar to that of ray-tracing, whereby each image point $s_1 = (x_1, y_1)$ of the SCI, the object coordinates (the X, Y, and Z coordinates defined in the local laser scanning system) of the corresponding 3D point can be computed. This database, termed the Point Cloud Visual Index (PCVI), allows computation of the exterior orientation of the RCI with respect to the point cloud by using simple image matching concepts, such as interest operators and correspondence matching (described in Chapter 4).

The assumption here is that the image correspondences are actually projections of the same 3D points of the point cloud. After interpolating the grey level of each pixel of the SCI from the return signal of the laser pulse, the same interpolation coefficients can be used to estimate the 3D coordinates for the centre of each pixel. That is, the results of the collinearity model (described in Section 3.3.1) and the interpolation (described in Section 3.3.2) are merged with the original point clouds to produce a database.

5.3.2.2 The Linking of Object Coordinates of the Point Cloud to the Pixels of the SCIs

This database consists of the following attributes for each pixel in the SCI: image coordinates (x, y), image coordinates (rows, columns), object space coordinates

(X, Y, Z) and the intensity values (I). Also, this index can be extended to include values such as interpolation constants and points to the original cloud.

The results of the correspondence matching (as described in Section 5.5) are a set of matched pixel values in column and rows. These values are queried using simple query language (SQL) programmed in Microsoft Access. This means that once matches are obtained between the SCI and RCIs, the object coordinates represented by the matched pixel are then available for use in the exterior orientation computation. This approach, which is a control point-free method, has important applications in terrestrial photogrammetric engineering (Styliadis *et al*, 2003). Also, solving the camera positions and orientations, the RCI can be reprojected onto the point cloud surface to produce a photorealistic model (discussed in Chapter 7).

5.3.3 Implementation of the Automatic Exterior Orientation

Here the measurement of image coordinates is implemented as an automated process. Experiments were done that used a Harris corner detector to extract candidates registration points in each image. Automatic image matching using ZNCC was performed on the RCI and SCI. The pixels intensity values of the SCI and the object coordinates were linked for use in the exterior orientation.

5.3.3.1 Initial Values of Exterior Orientation Parameters

As discussed in Chapter 2, the mathematical basis of the bundle adjustment is the collinearity model, and the collinearity equations are extended by the incorporation of parameters for self-calibration and the use of additional parameters that are supposed to model the systematic image errors (Fraser, 1997). The correction terms of the image coordinates are functions of these additional parameters. The initial values of the interior orientation parameters or internal camera parameters, such as a focal length and principal point coordinates, were given as the nominal values, and the lens distortion parameter was set to zero in this research. In addition, a set of correspondences in the images (image coordinates) with their corresponding control points (3D object coordinates of the point clouds) were used in the bundle adjustment computation.

Solving for the self calibrating bundle adjustment means estimating the additional parameters, as well as the exterior orientation and position of the cameras, the a priori standard errors of the exterior orientation and position, and a priori standard errors for both object and image coordinates. Camera-invariant additional parameters (APs) were used to model the systematic errors. These parameters were radial lens distortions (K1, K2), and decentring lens distortions (P1, P2) and principal point offsets (x_o, $y_{o)}$. All of the three different sets of data (A, B, and C) were processed with these a priori standard deviations.

5.3.3.2 Computation of Exterior Orientation Parameters

As discussed in the previous section, each pixel of the 2D image has the object point coordinates defined in the laser scanning system and in this section, the matches (image coordinates) and their corresponding object coordinates are used for the exterior orientation computation with the self-calibrating bundle adjustment approach. The bundle adjustment allows the simultaneous determination of the unknown object coordinates exterior orientation and interior orientation, with all relevant system parameters of the imaging system.

In addition, standard deviations are computed for all parameters, which give a measure of the quality of the imaging system. All input parameters are treated as unknown parameters (or as observed quantities) and they are input with their standard errors. Three sets of data each comprising four models were processed by the bundle adjustment. The solution to the adjustment was considered to been have achieved when the corrections to the camera positions and object point parameters and the image coordinate residuals were less than 50μm (a limit set in the SPGA program used).

5.3.4 Accuracy of the Bundle Adjustment

This section deals with the assessment of performance of the bundle adjustment based on the accuracy and precision of all reconstructed (or computed) parameters: 3D object points coordinates, image points coordinates, camera position and rotations.

The assessment also includes statistical analysis of the a posterior reference variance factor (discussed in Section 5.3.4.4) and the significance of the additional parameters - APs.

As mentioned before, for evaluation and comparison purposes, three sets of data each comprising for models were available for assessment. Data set A (representing a low complexity scene) contains simple continuous geometric features, data set B (representing a medium complexity scene) is of simple continuous geometric features with different objects, and data set C (a high complexity scene) contain many geometric features with different sizes and shapes. The results of the bundle adjustment of the three data sets are now presented

5.3.4.1 Object and Image Point Accuracy of Data Set A

The RMS errors (which are indications of accuracy) of the 3D object and image points coordinates for data set A are given in Table 5.3. For all the models for data set A, the largest RMS errors in X (MX) was 0.004m, in Y (MY) was 0.004m and in Z (MZ) was 0.007. The adjustment produced positional accuracy giving numerically better results compared to the nominal positional error (Se_{xy}) of ±6mm. The Z coordinate errors were smaller than ±8mm ($2Se_z$). The average accuracy in the object space coordinate determination for this dataset (MXYZ) was within 0.005m which was less than (Se_z).

Table 5.3: RMS Errors of Computed Object and Image Coordinates for Data Set A

Number of points = 68

RMS Errors						
	object coordinates (m)				image coordinates(mm)	
	MX	MY	MZ	MXYZ	Mx	My
Model_1	0.004	0.003	0.005	0.004	0.002	0.006
Model_2	0.005	0.004	0.007	0.005	0.001	0.006
Model_3	0.004	0.004	0.006	0.005	0.003	0.006
Model_4	0.005	0.004	0.006	0.005	0.001	0.006

Also, the maximum accuracy for image measurement for all models in x (Mx = 0.003mm) and y (My = 0.006mm) for data set A was less than one pixel (the pixel size of the images is 0.008mm). The notable results for data set A is the large size of the Y image coordinate errors (RMS error of 0.006mm for all models) with respect to the X image coordinate errors (which varies from 0.001mm to 0.003mm). This trend did not occur in other data sets and it therefore, requires further investigations for a more conclusive evaluation. The bundle adjustment produced a similar set of RMS errors for both X and Y object coordinates and the RMS Z coordinate error for all models was numerically larger than the RMS X and Y coordinate errors.

5.3.4.2 Object and Image Point Accuracy of Data Set B

The RMS errors of object and image point coordinates are shown in Table 5.4. For the model in dataset B, the RMS errors of objects point coordinates are numerically smaller than the RMS errors of object point coordinates for data set A (presented in the previous section). As with data set A, the RMS errors of object coordinates in X and Y are very similar (the bundle adjustment produces similar RMS errors) with the maximum accuracy in X being 0.004m and in Y being 0.004m. These were less than the Se_{xy} of ±6 mm.

Table 5.4: RMS Errors of Computed Object and Image Coordinates for Data Set B

Number of points = 50

RMS Errors						
	object coordinates (m)				**image coordinates(mm)**	
	MX	MY	MZ	MXYZ	Mx	My
Model_1	0.003	0.003	0.001	0.002	0.004	0.009
Model_2	0.004	0.002	0.003	0.003	0.001	0.005
Model_3	0.004	0.004	0.002	0.003	0.002	0.006
Model_4	0.004	0.002	0.003	0.003	0.003	0.006

However, the RMS Z coordinates error of the data set B are numerically smaller (with a maximum RMS error of 0.003m which was less Se_z of ±4mm) than the RMS Z coordinates errors (with a maximum RMS error of 0.007m) of data set A for all models in both data sets. The average accuracy in the object space coordinate determination for this dataset (MXYZ) was within 0.003m, which is also less than the average object point coordinate accuracy of data set A.

The maximum accuracy (Mx = 0.001mm and my = 0.006mm) for image measurement for all models, with the exception of model_1, in this data set was less than one pixel (8μm). The accuracy of image measurement for model_1 was 0.009m which was slightly more than one pixel.

5.3.4.3 Object and Image Point Accuracy of Data Set C

The object and image points accuracy, in terms of RMS values for data set C are presented in Table 5.5. It should be noted the maximum RMS errors for object point coordinates in X and Y are identical to the maximum RMS errors in X and Y for both data sets A and B. As with data set A, the RMS errors for Z object coordinates for all models for this data set is numerically larger than the RMS errors for all model for X and Y coordinates.

Table 5.5: RMS Errors of Computed Object and Image Coordinates for Data Set C

Number of points= 42

RMS Errors						
	object coordinates (m)				image coordinates(mm)	
	MX	MY	MZ	MXYZ	Mx	My
Model_1	0.001	0.002	0.007	0.004	0.004	0.005
Model_2	0.004	0.004	0.007	0.005	0.003	0.005
Model_3	0.002	0.002	0.007	0.004	0.002	0.004
Model_4	0.001	0.002	0.003	0.002	0.004	0.005

The RMS image coordinates for this data set is similar to the RMS image coordinates of the other two data sets previously discussed. The maximum RMS errors in both image coordinates x and y are less than one pixel.

5.3.4.4 The Reference Variance Factor

The reference variances for each data set for each mode are given in Table 5.6: These are highly depended on the a priori standard deviations. That is, the reference variances vary with changes in the a priori standard deviations. As can be seen from Table 5.6, the reference variances of the adjustment are generally similar for all the three data sets for all models. For model_1 of data set B has the highest value of reference variance factor.

Table 5.6: Standard Deviation of Unit Weight for all Data Sets

Standard Deviation of Unit Weight			
Data Set	**A**	**B**	**C**
Model_1	1.10	1.65	1.04
Model_2	1.30	0.93	1.07
Model_3	1.18	0.99	0.82
Model_4	1.13	1.12	1.03

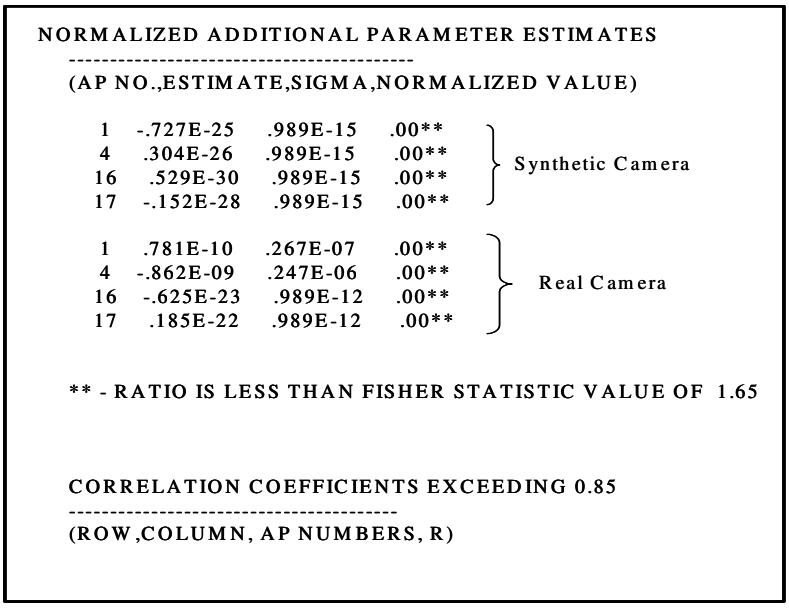
This value is statistically identical to that of the critical value at 5% level of significance. Data set C has reference variances for all models closer to unity than the reference variances for all model of data set A. This may mean that the a priori standard deviations for the camera parameters of data set C are more accurate estimations of their true values. Therefore, on average, the a priori standard deviations for all observations for all the datasets are correctly estimated.

5.3.4.5 Analysis of Additional Parameters Estimates

The APs were used in the BA process and these were checked for numerical stability, statistical significance and reliability in order to justify their presence and to avoid over-parameterization. The APs are usually correlated with one another or with camera parameters. Both the Fisher test and the correlation were used for accessing the significance of the APs on the BA solution.

Insignificant APs do not affect the solution of the bundle adjustment but may weaken the covariance matrix without positive contribution to the functional model.

A limit of 0.85 for the coefficient of the correlation was set as a threshold. If this limit is exceeded, the corresponding additional parameter will be excluded from the adjustment. This decision was based on the test values of the Fisher test which checks the significance of a single additional parameter. That is, if a "normalized" AP estimate was less than the assigned Fisher value, the AP is flagged as being not statistically significant at the chosen confidence level (Fraser, 1983). Figure 5.3 is an example of the results of the APs for both synthetic and real cameras.



NORMALIZED ADDITIONAL PARAMETER ESTIMATES

(AP NO.,ESTIMATE,SIGMA,NORMALIZED VALUE)

1	-.727E-25	.989E-15	.00**	Synthetic Camera
4	.304E-26	.989E-15	.00**	
16	.529E-30	.989E-15	.00**	
17	-.152E-28	.989E-15	.00**	
1	.781E-10	.267E-07	.00**	Real Camera
4	-.862E-09	.247E-06	.00**	
16	-.625E-23	.989E-12	.00**	
17	.185E-22	.989E-12	.00**	

** - RATIO IS LESS THAN FISHER STATISTIC VALUE OF 1.65

CORRELATION COEFFICIENTS EXCEEDING 0.85

(ROW,COLUMN, AP NUMBERS, R)

Figure 5.3: Example of APs for Synthetic and Real Cameras

Each of the APs is identified by the following numbers: 1 – radial distortion (K1), 4 – decentring distortion (P1), 16 – principal point offset in x (x_o), and 17 – principal point offset in y (y_o). The results indicate that APs were not significant in all the models for all the datasets confirming that the APs used neither under-constrained nor over-constrained the solutions obtained. This also indicates that the radial lens distortions, decentering distortions, and principal point offsets were insignificant on the solutions of the BA.

5.4 Discussion

For manual measurement, the achieved accuracy (in terms of RMS errors of the object point coordinates) for all models for both data sets A and B was 0.005m.

The average RMS error of the computed object points coordinates (with the automatic measurement) for data set A was 0.005m, 0.004m for data B, and for data C. it was 0.005m. Similarly, the average RMS errors of the automatic image measurement for data sets A, B and C were: 0.006mm, 0.009mm, and 0.005mm, respectively. These accuracy results obtained are compatible with the nominal accuracy offered by the manufacturer of the laser scanner.

Considering the nominal accuracy of the laser scanner, the average point spacing of scanning (5mm), manual pointing to the conjugates image points, 3D coordinate extraction from the point cloud, and that the RCI was not calibrated (assumed principal distance, principal point location and no lens distortion

correction), these results show that the concept of the synthetic camera image is a feasible method for the multisensor fusion.

Generally, the results indicate that there were no conceptual errors in the developed methodology (especially the concept of generating SCIs and the PCVI). Factors which could also influence the precision and accuracy are the interpolation errors (for the generation of the SCIs) and the use of only two images, but these were not investigated.

5.5 Summary

In this chapter, the process of creating point cloud visual index (PVCI) has been discussed. This index was created by linking the object coordinates of the point cloud to the pixels of the generated synthetic images (SCIs).

Using the PVCI facilitates the exterior orientation of the RCI to the point cloud by using simple image matching concepts such as interest operators and the correspondence matching. The chapter also discusses the implementation of the determination of the exterior orientation parameters by using both manual and automatic methods. The manual method, which includes, interior and exterior orientations uses 3DMapper software StereoMaker. The generated SCI has been modeled with the photographs (Real Camera Image - RCI) to develop a geometric relationship between the digital camera and laser scanner.

Accuracy assessment of the manual measurement was done by assessing the RMS errors of the object point coordinates., and it was found that the maximum RMS error for all the four models for the data A and B was 0.008m. The automatic implementation to determine the exterior orientation parameters with simultaneous bundle adjustment approach is also discussed. The accuracy assessment of the automatic measure was made by analyzing the RMS errors for both image and object space coordinates. For all models of the three data sets, the accuracy of the object coordinate was 0.007m and for the image coordinate measurement; the measurement error was within one pixel.

CHAPTER 6

EVALUATION OF TEST RESULTS

6.1 Introduction

Two methods for registering a terrestrial scanner and arbitrarily located and oriented cameras were presented in this thesis. In the first method manual digitizing and existing photogrammetric application software for the registration process were used. The second method involved the automation of the registration process. In the previous chapters, experiments and results of applying these methods to different sets of data were shown. Also, the accuracy and precision of the manual and the automatic determination of the exterior orientation parameters have been discussed.

In this chapter, a more comprehensive comparative evaluation of the two methods using two different sets of data is conducted. Thus, the results of the manual measurement are used to compare and evaluate the proposed automatic orientation strategy for the fusion of the multisensor data. Also the exterior orientation parameters of the generated Synthetic Camera Image are discussed and compared.

6.2 Analysis of the Perspective Centre of the Laser Scanner

The fundamental concept of the Synthetic Camera Image (SCI) was developed and used. The conceptual elements of generating the Synthetic Camera Images

been solved in the field of photogrammetry and in the process, it was however assumed that the perspective centre of the laser scanner were equal to zero. This section therefore analyzes the exterior orientation results from both the manual and automatic techniques. In this section the (six) exterior orientation parameters of each of the four images is given in the form of the location of the perspective centers in the laser scanner coordinate system in meters and the rotation angles

6.2.1 Manual Measurement with Photogrammetric Application Software

Tables 6.1 and 6.2 present the values of the exterior orientation of the measurements for data set A and data set B. Both scenes were acquired with the same laser scanner and digital camera.

Table 6.1 Exterior Orientation Parameters of Data set A using Photogrammetric Software (Positional unit: meters; Angular unit: degrees).

	SCI	*RCI_1R*	*RCI_1L*	*RCI_2L*	*RCI_2R*
X_o	-0.007	-0.752	0.792	1.678	-1.565
Y_o	-0.047	-0.006	-0.080	0.085	0.003
Z_o	-0.027	0.229	0.300	0.196	0.288
ω	0.047	2.387	2.189	2.024	3.222
φ	-0.029	-6.159	5.708	12.025	-15.020
κ	0.013	-0.366	-1.433	-3.487	2.427

The RMS errors for data set A were variable from 0.007m to 0.047m for camera position and 0.013˚ to 0.047˚ for camera rotation. Similarly, for data set B, these RMS errors varied from 0.005m to 0.012m for camera position and 0.005˚ to 0.076˚ for camera rotation. The results indicate that on average the exposure station coordinates of SCI of data set B are closer to zero than the exposure coordinates of SCI of data set A.

Also, it can be seen that for camera position for data set A, the X and Z values are closer to zero than the Y value with the X value being the closest to zero. But for camera rotation, phi and kappa values are closer to zero than the omega value with the kappa being the closest. These exterior orientation parameters of laser scanner do confirm the assumption already discussed in Section 3.3.1.

Table 6.2 Exterior Orientation Parameters of Data set B using Photogrammetric Software (Positional unit: meter; Angular unit: degrees).

	SCI	***RCI_1R***	***RCI_1L***	***RCI_2L***	***RCI_2R***
X_o	0.008	0.395	-0.934	-0.846	0.371
Y_o	-0.005	-0.029	-0.017	0.049	0.005
Z_o	-0.012	-1.816	-1.989	-2.275	-2.389
ω	-0.041	-2.657	-2.623	-2.937	-6.615
φ	0.057	-3.821	-3.556	-4.290	-1.869
κ	-0.076	-0.007	-0.097	0.436	-1.117

For data set B, the results are similar to that of dataset A except that the Y camera position is the closet to zero and omega value is smallest. Apart from the X position coordinate error of data set A the Y position coordinate error was greater than the $2Se_{xy}$ (nominal positional accuracy of ±12mm) and the Z coordinate error was also greater than the $2Se_z$ (nominal angular accuracy = ±8 mm). However, the values of the positional errors for data set B were less than the nominal positional accuracy (Se_{xy}) of ±6 mm with the Z coordinate error being greater than the nominal range accuracy (Se_z) of ±8mm.

These errors are mainly influenced by interpolation errors results (which may affect the quality of the generated synthetic images), systematic errors between terrestrial and photogrammetric measurements, and errors due to the measurements of the laser scanner coordinates themselves. However, the results indicate that on average the recovered exterior orientation parameters of laser scanner do confirm the assumption already discussed in Section 3.3.1 of Chapter 3.

6.2.2 Automatic Measurement with Bundle Adjustment

To verify the validity of the matching algorithm, three different data sets (data set A, B and C) were used and the results of these are presented in Tables 6.3, 6.5 and 6.7. Also, Tables 6.4, 6.6 and 6.8 provide the standard deviations for the measurements of camera position and rotation. Five camera stations were used on this project, all of which were at approximately the same elevation.

6.2.2.1 Exterior Orientation Parameters and their Standard Deviations of Data Set A

The exposure station coordinates produced by the bundle adjustment are generally closer to zero for data sets A and B than those camera station coordinates produced with the StereoMaker Software.

Table 6.3: Exterior Orientation Parameters of Data set A using the Bundle Adjustment (Positional unit: meters; Angular unit: degrees)

	SCI	*RCI_1R*	*RCI_1L*	*RCI_2L*	*RCI_2R*
X_o	-0.005	-0.747	0.786	1.670	-1.570
Y_o	-0.052	-0.009	-0.080	0.079	0.001
Z_o	-0.027	0.223	0.295	0.194	0.297
ω	0.048	2.383	2.183	2.029	3.230
φ	0.028	-6.160	5.711	12.035	-15.026
κ	0.016	-0.361	-1.436	-3.494	2.434

What is most evident in the results is that, for each camera station for data set A, the standard errors for both camera position and rotation (with the maximum standard deviation of 0.001m for position and 18" for the rotation) for the generated synthetic camera image were numerically less than the standard errors for the other four camera stations. These are expected as the SCI defines the object space coordinate system and the quality of the SCI certainly has a strong influence on the quality of the reconstruction results.

Table 6.4: Standard Deviation of Exterior Orientation Parameters for Data Set A

Standard Errors						
	Camera Position (m)			Camera Rotation(minutes)		
	σX_o	σY_o	σZ_o	$\sigma \omega$	$\sigma \varphi$	$\sigma \kappa$
SCI	0.001	0.000	0.001	0.1	0.3	0.1
RCI_1R	0.003	0.006	0.017	3.4	2.9	0.3
RCI_1L	0.007	0.001	0.028	1.3	0.7	0.0
RCI_2L	0.013	0.006	0.024	3.6	6.1	0.4
RCI_2R	0.033	0.005	0.022	2.4	0.7	0.2

As previously pointed out, the resampling techniques, nearest neighbour, has been used to generate SCI so the errors involved are due to interpolation errors, errors between terrestrial and photogrammetric measurements or errors due to the measurements of the laser scanner coordinates themselves.

Also, the maximum standard errors for all camera position for Z values are numerically larger than the standard errors for X and Y. For each of the models of the RCIs, the Z camera position standard error produced by the BA solution was greater than the ±8mm ($2Se_{xy}$ = ±8 mm) and for all models, unlike X coordinate errors of model_2L and model_2L, camera position standard errors in X and Y were less than the two times the nominal positional errors ($2Se_{xy}$) of the laser scanner of ±12mm.

It should be noted that in general, the standard rotation error in kappa ($\sigma\kappa$) show the least decrease in the standard error for the camera rotation. The maximum value of the standard rotation error in kappa was 18" and the minimum standard rotation errors both in omega ($\sigma\omega$) and phi ($\sigma\varphi$) for the RCIs was 42". For this data set there was no general trend of standard errors for the camera rotation.

6.2.2.2 Exterior Orientation Parameters and their Standard Deviations of Data Set B

As can be seen in these tables (i.e., Tables 6.3, 6.5 and 6.7), data set B produces camera position coordinates which are substantially smaller than the camera position coordinates for data sets A and C. Also, the camera rotation values of the SCI for dataset A were closer to zero that those rotation values for data sets B and C.

Table 6.5: Exterior Orientation Parameters of Data set B using the Bundle Adjustment (Positional unit: meters; Angular unit: degrees)

	SCI	*RCI_1R*	*RCI_1L*	*RCI_2L*	*RCI_2R*
X_o	0.002	0.397	-0.927	-0.852	0.372
Y_o	-0.002	-0.027	-0.021	0.045	0.007
Z_o	-0.005	-1.820	-1.989	-2.284	-2.390
ω	-0.048	-2.652	-2.617	-2.931	-6.614
φ	0.064	-3.826	-3.547	-4.290	-1.868
κ	-0.074	-0.010	-0.088	0.439	-1.108

The camera rotation values of data sets B and C are similar. It should be noted that for data set A, the X camera position was almost zero with the Y and Z camera position having substantially larger values. The phi angle for data set A was the closet to zero with the omega angle being the largest. The camera position values of data set B were all zeros but with generally substantial larger values for the camera rotation values. As in Section 6.2.1, the X and Y camera positional errors were less than Se_{xy} (the nominal positional error of ±6mm) and the Z coordinate error is less than $2Se_z$ (nominal range accuracy of ±8mm). The X coordinate error of data set was less the ±6mm (Se_{xy}) whereas the Y coordinate error of the same data set was greater than ±12mm ($2Se_{xy}$).

Table 6.6: Standard Deviation of Exterior Orientation Parameters for Data Set B

Standard Errors						
Camera Position (m)				**Camera Rotation(minutes)**		
	σ_{X_o}	σ_{Y_o}	σ_{Z_o}	σ_{ω}	σ_{φ}	σ_{κ}
SCI	0.002	0.001	0.002	0.5	0.6	0.6
RCI_1R	0.001	0.002	0.005	1.2	0.9	0.6
RCI_1L	0.005	0.005	0.012	2.9	1.7	0.4
RCI_2L	0.008	0.003	0.009	1.5	4.7	2.2
RCI_2R	0.001	0.004	0.001	2.0	0.1	0.1

The standard camera position and rotation errors for data set B (as can be seen in Table 6.6) and data set A were identical. However, the Z standard errors in data set B are numerically smaller than the Z standard errors in data set A. It can be seen that the maximum Z standard error in data set B is approximately half of the maximum Z standard error in data set A. The maximum RMS error values in camera position are about 0.008m in X and 0.005m in Y. These values were within the limit of the $2Se_{xy}$ (nominal positional error of ±12mm).

As pointed out in the previous section, the generated synthetic camera image shows the least standard error for both camera position and rotation, with the maximum standard error of camera position being 0.002m (less than Se_{xy} = ±6 mm) and that of the camera rotation being 36". As with data set A, the BA produced similar sets of standard camera rotation errors for data set B.

6.2.2.3 Exterior Orientation Parameters and their Standard Deviations of Data Set C

For data set C (as can be seen in Table 6.7), both the X coordinate and Z coordinate were greater than ±12mm ($2Se_{xy}$) and ±8mm ($2Se_z$) respectively. There was slight variation in the rotation values for data set B. As can be noticed the exposure coordinates of the SCI for data set C were slightly higher than for data sets A and B. However, these exterior orientation parameters of the laser scanner also do confirm the assumption already discussed in previous section.

Table 6.7: Exterior Orientation Parameters of Data set C using the Bundle Adjustment (Positional unit: meters; Angular unit: degrees)

	SCI	*RCI_1R*	*RCI_1L*	*RCI_2L*	*RCI_2R*
X_o	-0.098	0.173	0.175	0.174	0.180
Y_o	0.011	-0.074	0.065	0.048	0.031
Z_o	0.028	0.558	0.58	0.616	0.615
ω	-0.039	0.541	-0.568	-2.436	-0.698
φ	-0.063	10.864	6.859	7.256	5.044
κ	0.078	-1.808	-0.393	1.445	-3.135

Table 6.8: Standard Camera Position and Rotation Errors for Data Set C

Standard Errors						
	Camera Position (m)			**Camera Rotation(minutes)**		
	σX_o	σY_o	σZ_o	$\sigma\omega$	$\sigma\varphi$	$\sigma\kappa$
SCI	0.001	0	0.002	0	0.2	0.3
RCI_1R	0.002	0.002	0.013	0.3	1.0	0.4
RCI_1L	0.001	0.001	0.022	0.8	2.6	0.2
RCI_2L	0.000	0.003	0.000	0.6	0.9	0.6
RCI_2R	0.001	0.002	0.011	0.5	1.0	0.2

The standard deviations of the position and rotation measurements computed for data set C for all models Table 6.8. These are well within the accuracy

requirements discussed earlier. It can be seen that the omega (ω) and kappa (κ) rotational errors for all models were numerically smaller than the phi (φ) rotational errors. The maximum error for ω was 48", for κ it was 36", and 2.6 minutes for φ. As previously pointed out, the RCI was not calibrated and nearest neighbour resampling technique has been used to generate SCI so it is worth investigating the effects of resampling on the overall quality of the image and the impact of camera calibration. Further analysis of these effects is needed to better explain the high standard rotational errors for this data set.

6.3 Comparative Analysis of Manual and Automatic Matching results

Important for the multisensor image fusion is the correct determination of the orientation between 3D laser scanner and the 2D digital camera. The image orientations have been determined by classical least squares bundle adjustment and photogrammatric software using control points (defined in the laser scanner coordinate system).

The manual and automatic matching results have been examined to check the consistency of the image and object points coordinate measurements. The accuracy of these methods was assessed by the RMS errors. As was noted earlier, there was no known value upon which to test the relative orientation results against as there was no calibration data for either camera system.

This section, therefore, discusses the results of the direct comparison of the exterior orientation parameters determined by the photogrammetric software (manual) and the least squares bundle adjustment (automatic).

6.3.1 Using the Exterior Orientation Parameters

The differences between the measurements of the exterior orientation parameters for data sets A and B with the manual and automatic methods were computed and the results are given in Tables 6.9 and 6.10 for data set A and B, respectively. These are obtained by subtracting the automatic exterior orientation values from the manual orientation values for each model.

Table 6.9: Discrepancies of the Exterior Orientation Parameters with Data Set A (Positional unit: meters; Angular unit: degrees)

	SCI	*RCI_1R*	*RCI_1L*	*RCI_2L*	*RCI_2R*
δX_o	-0.002	0.005	0.006	0.008	0.005
δY_o	0.005	0.003	0.000	0.006	0.002
δZ_o	0.000	0.006	0.005	0.002	-0.009
$\delta \omega$	-0.001	0.004	0.006	-0.005	-0.008
$\delta \varphi$	0.001	0.001	-0.003	-0.010	0.006
$\delta \kappa$	-0.003	0.005	0.003	0.007	-0.007

Table 6.10: Discrepancies of the Exterior Orientation Parameters with Data Set B (Positional unit: meters; Angular unit: degrees).

	SCI	*RCI_1R*	*RCI_1L*	*RCI_2L*	*RCI_2R*
δX_o	0.006	-0.002	0.007	0.006	-0.001
δY_o	-0.003	-0.002	0004	0.004	-0.002
δZ_o	-0.007	0.004	0.000	-0.009	0.001
$\delta \omega$	0.007	0.005	-0.006	-0.006	-0.001
$\delta \varphi$	-0.007	0.005	-0.009	0.000	-0.001
$\delta \kappa$	-0.002	0.003	-0.009	-0.003	-0.009

Five camera stations were used in this thesis, all of which were approximately at the same elevation. This means that a positive value shows that the manual errors were larger than those of the automatic method. In the case of data set A, the discrepancy (difference between the manual and the automatic results) of the camera position (i.e. δx_o, δy_o, and δz_o) lies between 0.000 and 0.009m for all models and for camera orientation ($\delta\omega$, $\delta\varphi$, $\delta\kappa$) it lies between 0.001 and 0.010 degrees. The maximum discrepancy value (0.008m) for X and Y camera position was less ±12mm ($2Se_{xy}$) and that for Z coordinate (0.009m) was slightly greater than the $2Se_z$ (nominal range accuracy of ±8mm).

It can be seen from Table 6.9 that the discrepancy values for the exterior orientation of the generated synthetic camera image (SCI) were significantly smaller as compared to the other real camera stations.

The second data set (data set B) show similar results in case of the discrepancy of the exterior orientation parameters with discrepancy values for the camera position ranging between 0.000 and 0.009m and with the camera orientation ranging between 0.000 and 0.009 degrees. In both cases (i.e., the two methods used in the analysis), the computed parameters of exterior orientation parameters were similar (i.e., no substantial difference).

6.3.1 Using the Object Point Coordinates

To evaluate the differences in reconstructed object point coordinates of the bundle adjustment and StereoMaker, 10 points were used as checks for both measurements.

For each model in each data set, the check point coordinates were compared by, calculating the differences in X, Y, Z coordinates (x_dev, y_dev, z_dv). These represent the differences between the model coordinates as produced by manual and automatic pairwise measurements in the laser scanner coordinate system. The vectors of the discrepancies for all models for data sets A and B are graphed in Figures 6.1 and 6.2 respectively. The statistics (the minimum and maximum values and the RMS errors) of the discrepancies for individual models are presented in Tables 6.11 and 6.12 for data sets A and B, respectively.

Table 6.11: Statistics of Discrepancies of Object Point Coordinates for Data set A

Residuals(m)					
Coordinate Components		**Model_1**	**Model_2**	**Model_3**	**Model_4**
x_dev	**Min (m)**	0.000	0.000	0.000	0.000
	Max (m)	0.004	-0.005	0.005	0.005
	Mean (m)	0.002	0.002	0.002	0.003
	RMS(m)	0.001	0.002	0.002	0.002
y_dev	**Min (m)**	0.000	0.000	0.000	0.000
	Max (m)	-0.006	0.007	-0.007	0.006
	Mean (m)	0.003	0.003	0.002	0.002
	RMS(m)	0.002	0.002	0.002	0.002
z_dev	**Min (m)**	0.000	-0.001	-0.001	-0.001
	Max (m)	-0.006	-0.007	-0.008	0.008
	Mean (m)	0.003	0.004	0.004	0.003
	RMS(m)	0.002	0.003	0.003	0.002

What is most evident in the results is that, most of the models have RMS errors of approximately 3 mm (in X and Y and Z) which were less than ±6mm (Se_{xy}). For data set A, the RMS errors of the computed 3D coordinates difference were: 2mm [X], 3mm [Y], and 3mm [Z] and for data set B 3mm [X], 2mm [Y], and 2mm [Z]. In spite of the small discrepancies (i.e., with maximum RMS error of discrepancy of 3mm), conformance (i.e., agreement of positions measurement by manual with

photogrammetric software and computed positions by the automatic bundle adjusted) is generally very high and certainly demonstrates the accuracy and potential of the generated SCIs and the concepts of PVCI in the automation of the exterior orientation process.

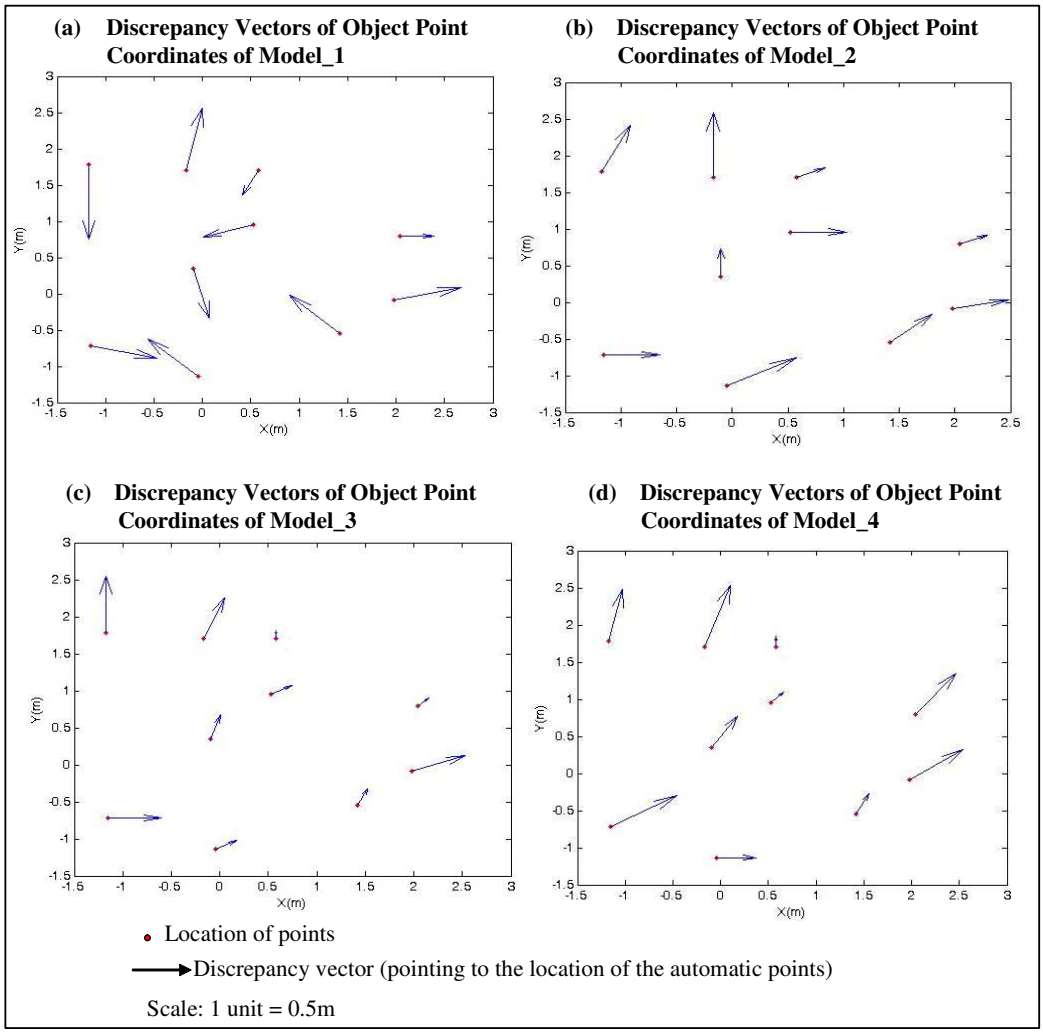


Figure 6.1 Discrepancy Vector of Object Point Coordinates for Data Set A (Deviations are in meters)

Table 6.12: Statistics of Discrepancies of Object Point Coordinates for Data Set B

Residuals(m)					
Coordinate Components		**Model_1**	**Model_2**	**Model_3**	**Model_4**
x_dev	**Min (m)**	0.000	0.000	-0.001	0.000
	Max (m)	0.006	0.007	0.009	0.007
	Mean (m)	0.002	0.003	0.003	0.003
	RMS(m)	0.002	0.003	0.003	0.003
y_dev	**Min (m)**	0.000	0.000	-0.001	0.000
	Max (m)	0.006	0.004	0.007	0.004
	Mean (m)	0.002	0.001	0.003	0.002
	RMS(m)	0.002	0.001	0.002	0.001
z_dev	**Min (m)**	0.000	0.000	0.000	0.000
	Max (m)	0.004	-0.006	-0.004	0.006
	Mean (m)	0.001	0.002	0.001	0.002
	RMS(m)	0.001	0.002	0.001	0.002

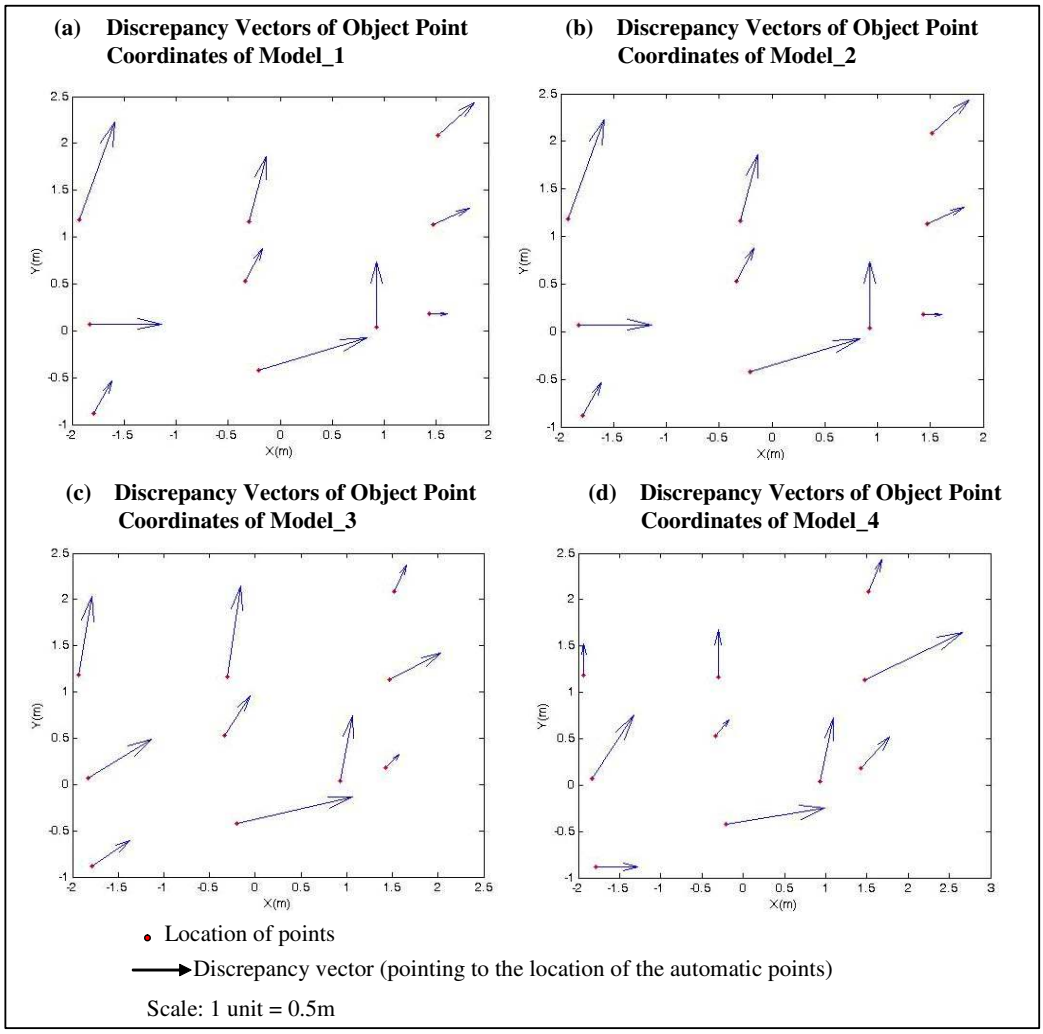


Figure 6.2 Discrepancy Vector of Object Point Coordinates for Data Set A (Deviations are in meters)

6.4 Summary of Results

The results of the adjustment were assessed using RMS errors of the object point and image coordinates. The average RMS errors of coordinates difference for all the models was 3mm for data set A and 2mm for data B. The residuals were very small, indicating highly precise networks had been established.

In addition, the object point coordinate and camera exposure values computed from the photogrammetric software agree closely with those resulting from the bundle adjustment where the orientation process was automated. Some high deviations values for some models might be due to systematic errors between terrestrial and photogrammetric measurements (such as manual pointing to the conjugates image points) and, also, the error due to the measurements of the laser scanner coordinates (such as 3D coordinate extraction from the point cloud) themselves

However, the results generally indicate that there is no significant difference between the quality of laser scanner measurements and those generated by the matching algorithm.

CHAPTER 7

APPLICATIONS OF THE DEVELOPED MULTISENSOR FUSION METHOD

7.1 Introduction

In Chapter 5, the object coordinates of the point clouds have been linked to the pixels of the generated synthetic camera images to create a visual reference (called point cloud visual index-PCVI). With the PVCI, it has been shown that automatic exterior orientation without the need for control point (it does not need retro-reflective or other special material-based targets) is a possibility. Also this control point-free method has important applications in terrestrial photogrammetric engineering. These applications include photorealistic 3D models presentation (that is texture-mapping the point cloud to create photo-realistic models), extraction of reference targets for registration and calibration purposes, automation of the registration of point clouds, and if the data are geo-referenced (the orientation of point cloud to an absolute coordinate system), they can be readily incorporated into existing GIS applications.

The solution to camera calibration (process of determining and relating the 3D position and orientation of the camera with respect to a certain world coordinate system) problems requires the knowledge of a set of corresponding 3D and 2D

features. Once matches are obtained between the SCIs and RCIs, the object coordinates represented by the matched pixel (i.e. to associate the 3D co-ordinates in point cloud with the corresponding pixel positions in the RCI) are then available for use in the exterior orientation computation. This solution (PVCI) which has been developed and implemented can be used for the automation of camera calibration.

Point cloud registration (combining several scans with different orientations and positions) is performed through the identification of common points in adjacent and overlapping scans (target-based registration). Once the points (at least three) are collected, a simple transformation can be determined. Using the PCVI would allow the automation of the process of establishing point correspondences between point clouds. Once matches are obtained between the SCIs (2D representation of a TLS point cloud) of point clouds, the object coordinates represented by the matched pixels are then immediately available for use in the point cloud registration process.

This following now describes the process of mapping the RCIs to the 3D geometric model to produce a geometrically correct, photorealistic view of the scene using the relative orientation and position parameters.

7.2 Mapping of Image Texture onto the Point Cloud

Texture mapping presented in this thesis is one of applications of the proposed multisensor fusion system. Texture-mapping the point cloud to create photo-realistic models which are essential for a variety of applications such as 3D city models, classification of real world objects, and virtual reality creation.

Texture mapping is basically a geometric transformation where pixels from the texture image are transformed to pixels of the surface. Then, for each pixel with this surface, the colour or grey value of the correspondent pixel in the texture image is assigned. This type of mapping has been used widely for over a decade to increase the sense of realism in computer generated scenes (Weinhaus and Deverajan, 1997).

Traditional texture mapping techniques, which simply involves the draping of static imagery over geometry, are not sufficient to represent highly detailed contents of the objects and scenes. In certain applications, it is crucial to preserve every line or edge and shape. The conceptual basis for texture mapping is depicted in Figure 7.1 (Weinhaus and Deverajan, 1997). In this Figure, there are two main transformations involved: One transformation maps the 3D objects surface and the computer screen space whilst the other transformation maps the 3D object surface and 2D texture space.

The discussions in this section focus on this latter transformation. The purpose of texture mapping is to produce a geometrically correct, photorealistic model of the scene and objects. There are numerous approaches for the photorealistic reconstruction of a general 3D scene or objects. The goal of all methods is the realistic rendering of the captured scene or objects (Stamos, 2001).

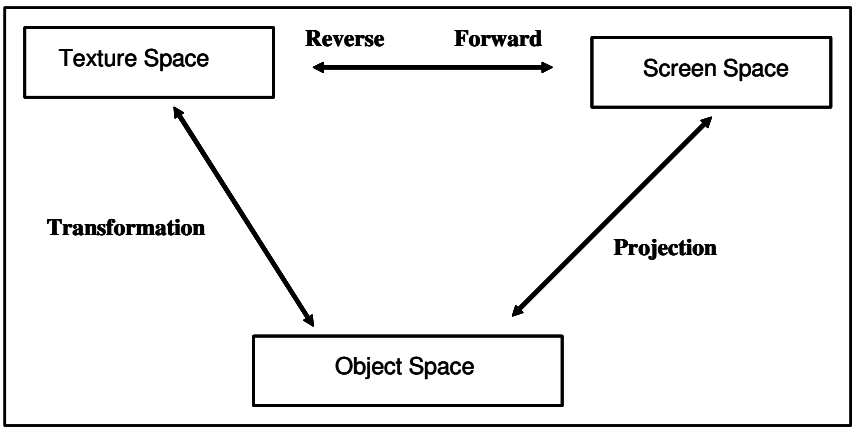


Figure 7.1: Texture Mapping Concept (after Weinhaus and Deverajan, 1997).

There are two major approaches in photorealistic texture mapping: geometric or range-based techniques such as laser scanners, and image-based approaches (Stamos, 2001). In the first category, a complete geometric model of the scene is required. Rendering from viewpoints can be computed by mapping images of the scene onto the geometric model (texture mapping). Most existing techniques for photorealistic texture mapping focus on this approach.

On the other hand, image-based modeling methods attempt to produce photorealistic renderings by skipping the geometric modeling steps. Set of images

of the scene generate a light field that is being interpolated in order to generate novel renderings from different view points. In this thesis, the range-based modeling is used.

7.2.1 The Process of Mapping

The geometry-based modeling approach applied in this research used laser scanning combined with high-resolution texture images. Thus, the reconstruction of a photorealistic texture model is achieved by the utilization of 2D imagery and 3D point cloud data. The 3D point cloud data is used for the building of the solid model, whereas the texture information is provided by an independent high resolution camera.

As previously mentioned, each pixel in the 2D photographic imagery is linked to its corresponding sampled 3D point on the object surface (Section 5.4.2.2). It should be noted that texture mapping described in this thesis is a view-dependent (i.e., a single scan and a single image of the scene) one. This form of texture-mapping is most effective when the model conforms reasonably closely to the actual structure of the scene (Debevec, 1996). In general, each image used is only a partial view of the scene or object. Thus, it is usually necessary to use multiple images in order to render the entire model of the scene.

In general, photorealitic texture mapping of 3D point clouds using 2D texture images consist of the following steps:

1. 3D point cloud and texture image acquisition of the from different viewpoints;

2. registration of images and the 3D point clouds into a common frame of reference;

3. Transformation of 3D point clouds into an intermediate surface-based representation; and

4. Merging of the surface-based representation and the texture image into a common representation.

The acquisition of the 3D point clouds and 2D texture image has been described in Chapter 3. The relative position of the range sensor and the camera is then known from the exterior orientation procedure (Chapter 5). The result of this procedure is that the intensity image and the geometric model are positioned and oriented in the same coordinate systems. With the position and the orientation of the 2D camera relative to the laser scanner known, the 2D photographic image is reprojected to the laser scanner frame of reference to create a texture map. Figure 7.4 shows a detailed 3D model, which was created.

A two-step approach is taken to achieve the realistic 3D model using the texture mapping of images and the 3D point cloud. As the first step, an accurate 3D model of the scene is obtained (Figure 7.4) using the mesh (Figure 7.3) generated in the Cyclone Software. In order to capture the complete geometric data of a

scene or object, the scene needs to be scanned from a number of different viewing directions to produce a set of complex and overlapping 3D point clouds. Generally, 3D modeling systems are based either on volumetric analysis or mesh-based (Stamos, 2001) and in this thesis the latter approach is adopted. In the mesh-based approaches, the 3D point clouds is transformed into a mesh of triangular faces.

In the second step, the texture image (i.e., the real camera image in Figure 7.5 is mapped onto the developed geometric model in Figure 7.6. This Figure presents how view-dependent texture mapping can be used to produce photorealistic models. This example uses only one data set (i.e., data set B)

7.2.2 Results and Analysis

Figure 7.6 represents the results of the photorealistic texture mapping process.

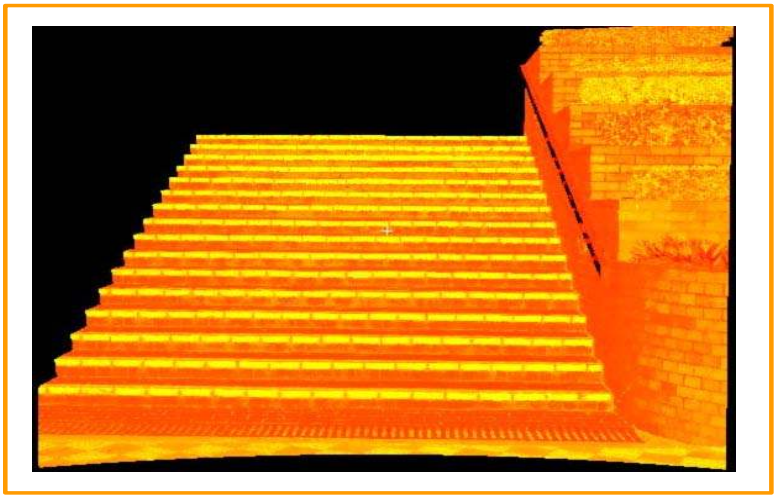


Figure 7.2: Raw Point Cloud

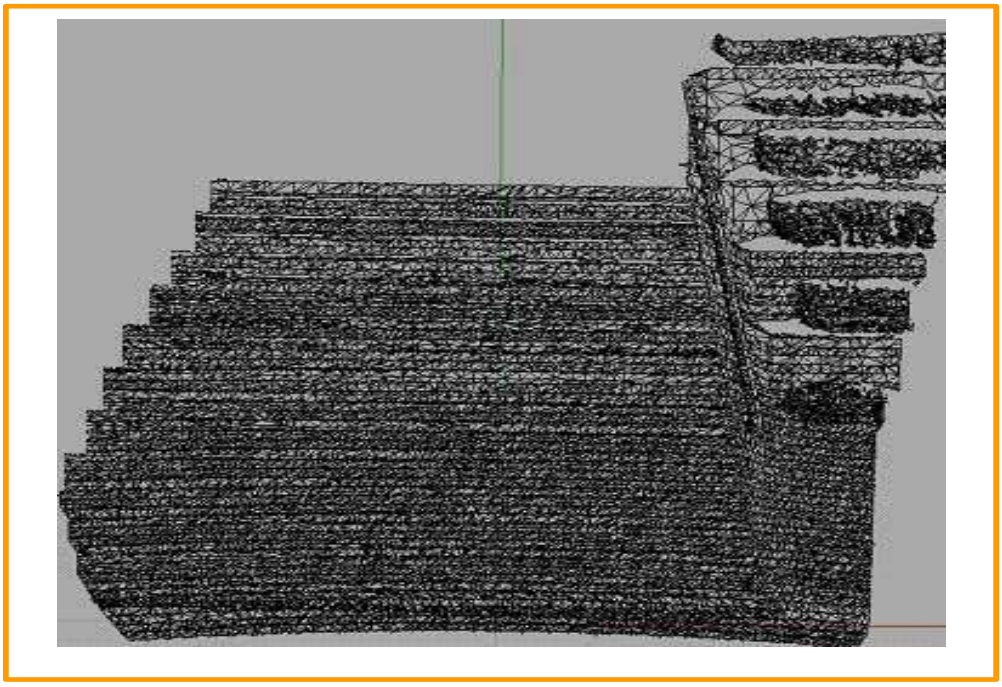


Figure 7.3: Snapshot of Simplified Mesh

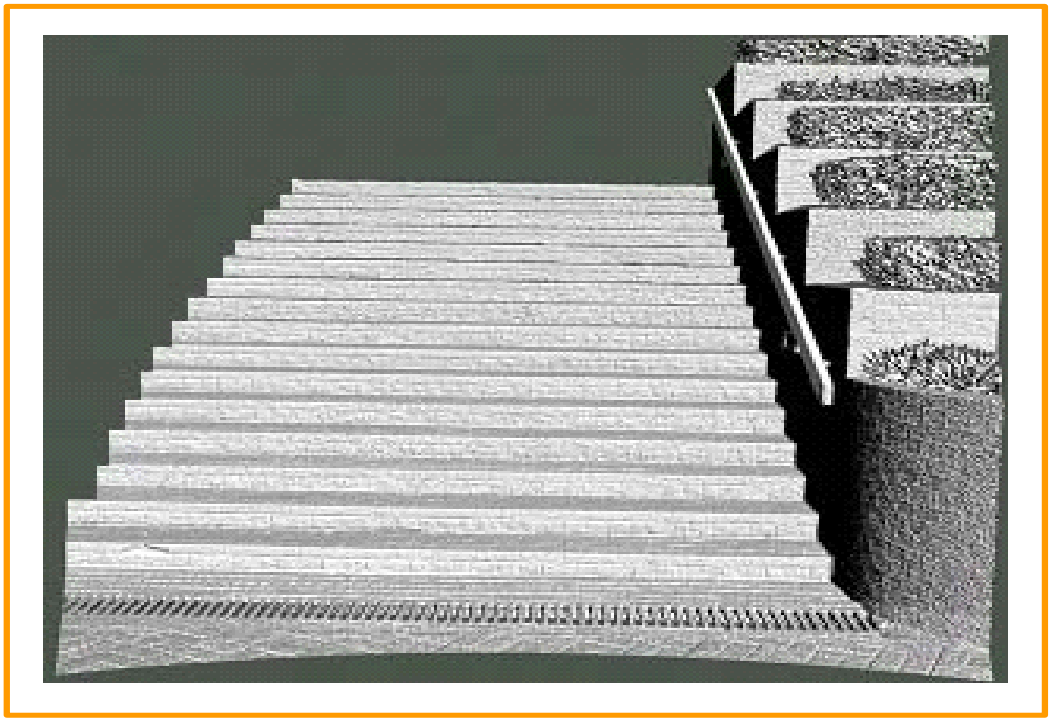


Figure 7.4: Geometric Model

Figure 7.5: Texture Image

Figure 7.6: Example of View-Dependent Texture Mapping

Mapping texture from multiple camera views onto a 3D model of a real object always involves potential problems. The main problems involve textural distortion at the boundaries of triangles, invisible surface parts of the real object, and highlights owing to inhomogeneous environment lighting. Since photo-realism is of utmost importance, it became apparent that many problems related to geometric and radiometric distortions had to be solved. In preserving every line or edge, even slightly curved surfaces pose problems of distortions and discontinuity.

Line and edge textures on the models can be visibly affected by any small geometric errors. Sources of these errors can be due to the mapping between triangle plane and image plane, and improper camera calibration and image registration. Since this project focuses on high resolution textures and high quality geometry, the texture size can end up being quite large, with memory limitations becoming a major issue. Another issue is mapping between the image and the developed model becomes easy and accurate if they overlap, they have the same resolution. However, mapping becomes difficult if the images contain unmodeled detail.

7.3 Summary

In this chapter, the potential applications of the proposed multisensor fusion method have been discussed. The proposed methodology allowed the automatic registration (the position and the orientation of the 2D camera relative to the laser scanner known) of point cloud and real camera images.

The registered real camera images were used as texture maps to enhance the 3D models generated from the 3D point cloud. The process of this mapping was described and the results presented.

CHAPTER 8

CONCLUSIONS AND RECOMMENDATIONS

8.1 Summary

The aim of this thesis was to develop a methodology to fuse data from high-density 3D point cloud data (such as those produced by terrestrial laser scanning systems-3D sensors) with high-resolution perspective imagery (such as that generated by digital cameras or scanned photographs- 2D sensors) and to produce high accuracy (geometric and radiometric) products such as photorealistic models.

The following sections summarize the data acquisition systems, generation of the synthetic camera images, registration methods and view-dependent texture mapping. In addition, conclusions, limitations of the study and possible related future work are discussed.

The strategy for terrestrial measurements (photogrammetry and laser scanning) have been discussed and the data sets used in this research have also been presented. In order to test both the fidelity of the SCI concepts and the correspondence algorithms, three deferent scenes were scanned with the laser scanner and for each scan, a series of 2D photographic image (RCI) were taken with a digital camera.

The photogrammetric collinearity model was used to fuse data from the two sensors (digital camera and terrestrial laser scanner) to generate synthetic images. This model was used to fuse data from the two sensors (digital camera and terrestrial laser scanner) to generate synthetic camera images equivalent to that of a digital camera. These images were created by interpolating the irregularly spaced synthetic data onto a grid at an even spacing. The generated synthetic images and the 2D photographic images were used in the registration process.

Initial stages of the research focused on undertaking the registration process using manual methods and existing photogrammetric (i.e.3DMapper software StereoMaker) application software. Accuracy assessment of the manual measurement was done by assessing the RMS errors of the object point coordinates.

The second phase involved the automation of image matching process. A Harris corner detector with appropriate threshold value was first used to extract feature points in both images. The extracted feature points were then matched with the zero-mean normalized cross-correlation (ZNCC) similarity metrics to find correspondences between the SCI and RCI. The algorithm computes similarity scores for every pixel in the image by taking a fixed window. A validity check was implemented by estimating correlation twice by reversing the roles of two images. The RANSAC algorithm has been used to filter out false (or incorrect) correspondences.

The process of creating a point cloud visual index (PCVI) has been discussed. This index was created by linking the object coordinates of the point cloud to the pixels of the generated synthetic images (SCIs). Using the PCVI allowed the exterior orientation of the RCI to the point cloud by using simple image matching concepts such as interest operators and the correspondence matching.

The automatic implementation to determine the exterior orientation parameters with simultaneous bundle adjustment approach is also discussed. The accuracy assessment of the automatic measure was made by analyzing the RMS errors for both image and object space coordinates and the exterior orientation parameters. With the position and the orientation of the 2D camera relative to the laser scanner known, the 2D photographic image was reprojected onto the laser scanner frame of reference to create a photorealistic texture map.

8.2 Conclusions

TLS and the digital camera capture their data with similar imaging geometry-perspective projection. This fact has been exploited to create synthetic camera images by integrating data from the two sensors. The fundamental conceptual elements of creating synthetic camera images have solved and developed in the field of photogrammetry. The generated SCI has been modeled with the photographs (Real Camera Image, RCI) to develop a geometric relationship between the digital camera and laser scanner.

The manual measurement process was necessary to understand the key issues such as geometric quality, both spatial and geometric resolutions of the generated synthetic camera image. Features were detected, extracted and matched to develop geometric relationship between the digital camera and laser scanner. Using the RMS errors object point coordinates as accuracy for the manual measurements, it was found that the maximum RMS error for all the four models for the data A and B was 0.008m.

The results generally indicated that the coordinates and the rotation angles of the perspective centre of the laser scanner were very close to zero (these exterior orientation parameters of laser scanner do confirm the assumption already discussed in Section 3.3.1). These results were supported and confirmed by the bundle adjustment which was used to reconstruct the 3D object space coordinates and to recover camera positions. The accuracy of the object coordinate determination is with 0.007m and for the image coordinate measurement error is within one pixel.

The RMS residuals for all models for all the three data sets were very small (RMS errors for all models of the three data sets was 0.007m for object coordinates and within one pixel (0.008mm) for image coordinates), indicating a precise network had been established. In addition, the object point coordinate and camera exposure values computed from the photogrammetric software agree closely with those resulting from the bundle adjustment where the orientation process was

automated. Some high deviation values for some models might be due to systematic errors between terrestrial and photogrammetric measurements and also, the error due to the estimation of the laser scanner coordinates themselves.

However, the results generally indicate that there is no detectable difference between the quality of laser scanner measurements and those generated by the manual and automatic matching algorithms. These results indicate that there were no conceptual errors in the developed methods.

Taking all models (i.e., image pairs) for all data sets into consideration, control point object coordinates and camera exposure coordinates show good accuracy (RMS errors for all models of the three data sets was 0.007m for object coordinates and within one pixel (0.008mm) for image coordinate), and indicate a good project solution. The results show that the concept of the synthetic camera image is a feasible method for multisensor data fusion.

Factors, which could also influence the precision and accuracy, are the algorithm robustness for matching or for the measurements of parallaxes, the influence of distortion parameters, and the use of only two images but these are not investigated.

8.3 Limitations of this Study

- This research focused on fusing data from close-range camera and terrestrial laser scanner. The performance of the proposed methodology is limited to data sets from these sensors. The developed concepts have not been tested with aerial photography and point clouds from airborne laser scanning systems (ALS) which have intrinsically much simpler geometric relationships.

- Although the Cyrax 2500 system can operate over distances of up to 100 metres, the data sets were effectively captured over a shorter range. That is the data sets used in this thesis was captured at a very close-range (<50 m range). Both geometry and reflections (energy reflected by the measured points) are key elements in developing the synthetic camera images for use in the matching process. Since reflectivity value of the scanner depends basically on the type of material that makes up the object and the range distance, measuring distances to materials with a low reflectivity at a long range (distances beyond 50m) may affect the quality of the data capture with some areas missing data.

As already pointed out, one of the salient properties of a scanner that strongly influences both point cloud resolution and positional uncertainty is the laser beamwidth so scanning at distances beyond 50m may increase the positional uncertainty of the laser beamwidth and may also result in low (coarse) vertical and horizontal scan density (i.e., low density sampling interval). In this case, more advance resampling techniques such as bilinear, bi-cubic and natural neigbour

must be used to generate the SCI. The effects of scanning at the full range (100m or over) of the scanner and the effects material of the object used on the quality of the generated SCIs are not investigated.

- The setup for data capture allowed a single scan with multiple images. That is the object geometry is first scanned and then several photographic images are taken to the left and to the right of the laser scanner, each in different viewing positions.

- The developed methodology did not include how to manage multiple point clouds and multiple images. Since each scan has its own reference system multiple scans requires the registration scans in a single coordinate system before generating the synthetic images. Also, the research did not investigate how far the developed methodology could be extended to cater for large differences in location and orientation between the camera and the laser scanner.

- The texture-mapping approach was a view dependent one. The current method did not merge multiples images on the 3D model. The method adopted only texture-mapped one image per view.

8.4 Recommendations for Future Work

Research has shown that, using simple techniques, combining both types of data – photographs (Real Camera Image, RCI) and point clouds - is feasible and can lead

to a very flexible system for capturing 3 dimensional information about an object. This research highlighted several areas that, in order to fully exploit the positive characteristics of the two technologies, need to be further investigated, developed and refined. Upon resolving these issues, the potential for integrating TLS point clouds and independently taken photographs, both for their image quality and photogrammetric potential, will provide a significant boost and will allow both these technologies to be used in an integrated environment greatly enhancing the usefulness and efficiency of the measurement and modeling processes. In future investigations the followings points should be further addressed:

- Investigation into the resolutions of the SCI. There are two options related to the resampling the backprojected laser point. The first option is, to generate the SCI by keeping the original resolution of the point cloud data (which is 1000 x 1000 at full scan) and then compute the pixel size. This means that the generated SCI will always be a square image of 1000 X 1000 pixels. The second option, is to keep the pixel size of the real camera image involved in the project (for the example, the Nikon D1x has 8 micrometer pixel size) and then compute the number of pixels or the resolution. In this case, the resolution of the generated synthetic image may vary depending on the maximum and minimum values in the datasets. This latter option was implemented and it is worth investigating the effect of the former on the overall quality of the image.

- Study the effects of the resampling techniques. Another issue involved with the quality of the SCI is the resampling techniques. The simplest of the resampling techniques, nearest neighbour, has been used to generate the SCI and the question is "could the more advance resampling techniques such as bilinear, bi-cubic and natural neighbour improve the quality and content of the SCI".

- Development of sub-pixel corner detectors algorithms and investigation into the effects of the combined use of edge and corner detection approach. The next aspect relates to the feature detection and the correspondence matching and what needs to be addressed is the impact of sub-pixel feature detection and the use of a combined edge and corner detectors approach instead of corners only on the overall matching results.

- To assess the impact that camera calibration on the matching results.

- Multiple scans and multiple images. This would also include how to manage multiple point clouds and multiple images.

- The effects of different base-height ratios on achieved accuracy.

LIST OF REFERENCES

3D Mapper, 2004. StereoMaker Description at http://wwww.3dmapper.com. Accessed, October 2004.

Adams, J. C., Smith, M. J., and Bingley, R. M., 2001. "Terrestrial Laser Scanning and Close-range Photogrammetry for Cliff-face Mapping". In Proceeding of Remote Sensing and Photogrammetry Society 1st Annual Meeting, pp. 537-548, London.

Antonio, J., Centeno, S. J. A. S., and Miqueles, M. A., 2004. "Extraction of Buildings in Brasilian Urban Environments Using High Resolution Remote Sensing Imagery and Laser Scanner Data". Proceedings, XXth ISPRS Congress, Istanbul, Turkey, 12-23 July, paper 589.

Armangue, X., Salvi, J., 2003. "Overall View Regarding Fundamental Matrix Estimation". Image and Vision Computing, 21 (2003), pp. 205–220.

Atkinson, K. B. (Ed)., 1996. *Close-range Photogrammetry and Machine Vision*, Bristol -1996.Wittles Publishing.

Baltsavias, E. P., Etienne, F., Bauder, A., Bocsch, H., and Pateraki, M., 2001. "Digital Surface Modelling by Airborne Laser Scanning and Digital Photogrammetry for Glacier Monitoring". The Photogrammetric Record. Vol. XVII, No.98, pp. 243-273.

Barbarella, M., Lenzi, V., and Zanni, M., 2004. "Integration of Airborne Laser Data and High Resolution Satellite Images Over Landslides Risk Areas". Proceedings, XXth ISPRS Congress, Istanbul, Turkey, 12-23 July , paper 945.

Barber, D., Mills, J., and Bryan, P., 2001. “Laser Scanning and Photogrammetry-21st Century Metrology”. CIPA 2001 International Symposium, University of Postdam, Germany, September, pp 18-21.

Behan, A., 2000. On the Matching Accuracy of Rasterised Scanning Laser Altimeter Data”. IAPRS, Vol. XXXIII, Amsterdam. 15p.

Bennamoun, M., and Mamic, G. J., 2002. *Object Recognition: Fundamentals and Case Studies*, Sprinter, London.

Beraldin, J. A., Blais, F., Boulanger, P., Cournoyer, L., Domey, J., EL-Hakim, S. F., Godin, G., Rioux, M., and Taylor, J., 2000. “Real World Modelling through High Resolution Digital 3D Imaging of Objects and Structures”. ISPRS Journal of Photogrammetry and Remote Sensing. Vol. 56, No.4, pp. 230-250.

Bilodeau, G.-A., and Bergevin, R., 2000. “Geometric Modelling of 3D Objects from Single 2D Images”. 15th International Conference on Pattern Recognition, pp. 770-773.

Boehler, W., Bordas V. M., and Marbs, A., 2003. “Investigating Laser Scanner Accuracy”. In: IAPRS, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, Part 5/C15, Antalya, Turkey, 30 Sept.4 Oct., pp 696-702.

Boughorbal, F., Page, D., and Abidi, M., 2000. “Automatic Reconstruction of Large 3D Models of Real Environments from Unregistered Data-sets”. SPIE Conference on 3D Image Capture and Applications III, Vol. 3958, pp. 234-243.

Brown, L.G., 1992. “A Survey of Image Registration Techniques”. ACM Computing Surveys, Vol. 24, No.4, pp. 325-376.

Bruder, S., Farooq, M., and Bayoumi, M., 1994. “A Feature-Level Approach to Multisensor Integration; Emphasizing Robotic Applications”. IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems, Las Vegas, Nevada, pp. 477-484.

Canisius, F. X .J., and Turral, H., 2003. “Fusion Technique to Extract Detail Information from Moderate Resolution Data for Global Scale Image Map Production”. Proceedings, the 30th International Symposium on Remote Sensing of Environment – Information for Risk Management and Sustainable Development – November 10-14, Honolulu, Hawaii.

Capel, D., and Zisserman, A., 2003. “Computer Vision Applied to Super-Resolution”. IEEE Signal Processing Magazine. pp. 75-86.
http://www.robots.ox.ac.uk/~vgg/publications/papers/capel03.pdf.

Chandler, J. H., Shiono, K., Rameshwaren, P., and Lane, S. N., 2001. “Measuring Flume Surfaces for Hydraulics Research Using KODAK DCS460”. The Photogrammetric Record. Vol. XVII, No.97, pp. 39-61.

Chen, Q., Defrise, M., and Deconinck, F., 1994. “Symmetric Phase-only Matched Filtering of Fourier-Mellin Transform for Image Registration and Recognition”. IEEE Transaction on Pattern Analysis and Machine Intelligence, pp 1156-1168.

Coorg, S., and Teller, S., 1999. “Extracting Textured Vertical Facades from Contolled Close-range Imagery”. In IEEE Conference of Computer Vision and Pattern Recognition, Fort Collins, Colorado, pp 625-632.

Cramer, M., 2002. “Investigations on Long Term Stability of System Calibration for Direct Georeferencing”. Final Project Study, Institute for Photogrammetry (ifp), University of Stuttgart, May, 38 pages.

Crowley, J. L., and Martin, J., 1995. “Experimental Comparison of Correlation Techniques”, IAS-4, International Conference on Intelligent Autonomous Systems, Karlsruhe. http://www-prima.imag.fr/Prima/Homepages/jlc/papers /IAS95-martin.pdf.

Csatho, B., and Schenk, T., 1998. “Multisensor Data Fusion for Automatic Scene Interpolation”. In International Archives of Photogrammetry and Remote Sensing. Vol. XXXII (7), pp. 336-341.

Cyrax, 2005. Leica Geosystems – Cyrax Product Specification at http://www.cyra.com/products/HDS3000_specs.html (Link checked, November 2004).

Dai, X., and Khorram, S., 1999. “A Feature-Based Image Registration Algorithm Using Improved Chain-Code Representation Combined with Invariant Moments”. IEEE Transactions on Geoscience and Remote Sensing, Vol. 37, No. 5, pp. 2351-2363.

Dakowicz, M., and Gold, C., 2003. “Extracting Meaningful Slopes from Terrain Contours”. International Journal of Computational Geometry and Applications, Vol. 13, No. 4, pp. 339-357.

Dare P., and Dowman I., 2001. “An Improved Model for Automatic Feature-Based Registration of SAR and SPOT Images”. ISPRS Journal of Photogrammetry and Remote Sensing. Vol. 56, No.1, pp. 13-28.

Dare P., and Dowman I., 2000. “A New Approach to Automatic Feature Based Registration of SAR and SPOT Images”. International Archives of Photogrammetry and Remote Sensing. Vol. XXXIII, Part B2, pp. 125-130.

Debevec, P. E., 1996. “Modeling and Rendering Architecture from Photographs”. PhD Thesis, Computer Science Division, University of California, Berkeley, pp 305.

De Castro, E., and Morandi, C., 1987. “Registration of Translated and Rotated Images using Finite Fourier Transforms”. IEEE Transactions on Pattern Analysis Machine Intelligence, Vol. PAMI-95, pp. 700-703.

Dias, P., Sequeira, V., Goncalves, J.G.M., and Vaz, F., 2002. “Automatic Registration of Laser Reflectance and Colour Intensity Images for 3D Reconstruction”. Robotic and Autonomous Systems, No. 39, pp. 157-168.

Dufournaud, Y., Schmid, C., and Horaud, R., 2004. “Image Matching Scale Adjustment”. Computer Vision and Image Understanding, 93 (2), February 2004. pp. 175-194.

El-Hakim, S. F., and Beraldin, J.-A., 1994. ”On the Integration of Range and Intensity Data to Improve Vision-Based Three-Dimensional Measurement”. In Videometric III, Proceeding of SPIE, Vol.2350, pp. 306-327.

Elstrom, M. D., Smith, P.W., and Abidi, M.A., 1998. “Stereo-Based Registration of LIDAR and Colour Imagery”. SPIE Conference on Intelligent Robots and Computer Vision XVII: Algorithms, Techniques, and Active Vision, 3522, pp. 343-354.

Essadiki, M., 2004. **“**New Technique for Combining Panchromatic and Multispectral Spot Images for Multipurpose Image-maps”**.** Proceedings, XXth ISPRS Congress, Istanbul, Turkey, 12-23 July, paper 748.

Fischler, M. A., and Bolles, R. C., 1981. “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography”. Comm. ACM, 24(6), pp. 381–395.

Forkuo, E. K., and King, B., 2004a. "Registration of Photogrammetric Imagery and Laser Scanner Point Clouds". Proceedings, Mountains of Data, Peak Decisions, 2004 ASPRS Annual Conference, Denver, Colourado, 23-28 May, paper 0058.

Forkuo, E. K., and King, B., 2004b. "Automatic Fusion of Photogrammetric Imagery and Laser Scanner Point Clouds". Proceedings, XXth ISPRS Congress, Istanbul, Turkey, 12-23 July, paper 475.

Forkuo, E. K., and King, B. A., 2003. “Fusion of Photogrammetric Imagery and Laser Scanner Point Clouds”. Proceedings, 7th South East Asian Survey Congress, Hong Kong, 3-7 November, paper 49.

Fraser, C. S., 1997. Digital Camera Self-calibration, JPRS, April, Vol. 52, No. 4, pp. 149-159.

Fraser, C. S., 1983. SPGA Program Description, University of Calgary.

Fritscher, K . D., 2004. “A Software Framework for Pre-processing and Level-set Segmentation of Medical Image Data”. MSc. Thesis, Institute for Biomedical Image Analysis, University for Health Science, Medical Informatic, Austria, 85p.

Gamba, P., Dell’acqua, F., and Houshmand, B., 2003. “Comparison and Fusion of LIDAR and InSAR Digital Elevation Models Over Urban Areas”. International Journal of Remote Sensing, pp. 1–12.

Ghanbari, E., and Azizi, A., 2003. “A Rapid and Low Cost Photogrammetric System”. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, Part 5/W12. http://www.commission5.isprs.org/wg4/workshop_ancona/proceedings/35.pdf.

Ghassemian, H., 2003. “A Retina Based Multi-Resolution Image Fusion”, Proceeding of ISPRS2003, Challenges in Geospatial Analysis, Integration and Visualization, Sept. http://www.fzg.uni- osnabrueck.de/mitarbeiter/schiewe /papers/10.pdf.

Giachetti, A., 2000. “Matching Techniques to Compute Image Motion.” Image and Vision Computing, 18, pp. 247-260.

Gonzalez, R., and Woods, R., 2002. *Digital Image Processing.* Addison-Wesley, 2nd edition.

Gordon, S., Lichti, D., and Stewaart., M., 2001a. “Application of a High-resolutiom Ground-based Laser Scanner for Deformation Mearsurements”. The 10th FIG Internal Symposium on Deformation Measurements, 19-22 March, California , USA, pp. 24-32.

Gordon, S., Lichti, D., Stewart, M., and Taskirie, M., 2001b. “Metric Performance of a High-resolution Laser Scanner”. SPIE Conference on Videometrics and Optical Methods for 3D Shape Measurements, SPIE Proceedings, Vol.4309, pp. 174-183.

Glover, D., Jenkins, B., and Doney; S., 2004. “Measurement Theory, Probability Distributions, Error Propagation and Analysis”. On line Lecture Notes. http://w3eos.whoi.edu/12.747/chaps/ (Accessed on 10 June 2005).

Guarnieri, A., Remondino, F., and Vettore, A., 2004. “Photogrammetry and Ground-based Laser Scanning Assessment of Metric Accuracy of the 3D Model of Pozzovehhiani Church”. FIG Working Week 2004. TS on "Positioning and Measurement Technologies and Practices II - Laser Scanning and Photogrammetry". Athens, Greece, 22-27 May.
http://www.photogrammetry.ethz.ch/general/persons/fabio/fig_atene.pdf.

Guidi, G., Tucci, G., Beraldin, J.-A., Ciofi, S .,Ostuni, D., Costantini, F., and. El-Hakim, S., 2002. “Multiscale Archeological Survey Based on the Integration of 3D Scanning and Photogrammerty”. Proceedings, International Workshop on Scanning for Cultural Heritage Recording-Complementing or Replacing Photogrammetry, Greece , 1-2 September, pp. 54-58.

Gunturk, B. K., 2003. “Multi-Frame Information Fusion for Image and Video Enhancement”. PhD Thesis, School of Electrical and Computer Engineering, Georgia Institute of Technology.

Guo, T., 2003. “3D City Modeling Using High-resolution Satellite Image and Airborne Laser Scanning Data”. PhD Thesis, Department of Civil Engineering, University of Tokyo, Tokyo.

Haala, N.; and Brenner, C., 1999. “Extraction of Buildings and Trees in Urban Environments”. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 54, Issue. 2, pp. 130-137.

Habib, A. F., and Alruzouq, R. I., 2004. “Line-based Modified Iterative Hough Transform for Automatic Registration of Multi-source Imagery”. The Photogrammetric Record, 19(105), pp. 5-21.

Harris, C. and Stephens, M., 1988. “A combined Corner and Edge Detector”. Proceedings of the 4th ALVEY Vision Conference, University of Manchester, England, September, pp. 147-151.

Hartley, R., and Zisserman, A., 2000. *Multiple View Geometry in Computer Vision*. Cambridge University Press, London.

Heipke ,C., Mayr, W., Wiedmann ,C., and Ebner, H., 1997. “Automatic Aerotriangulation with Frame and Three-line Imagery”. In Proceeding of SPIE Conference on Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision III. Vol.3072, pp 286-296.

Heipke, C., 1996. “Automation of Interior, Relative and Absolute Orientation”. International Archives of Photogrammetry and Remote Sensing, volume XXXI, Part B3, pp. 297-311.

Hong, G., and Zhang, Y., 2004. “The Effects of Different Types of Wavelets on Image Fusion”**.** Proceedings, XXth ISPRS Congress, Istanbul, Turkey, 12-23 July, paper 915.

Hsieh, J. W.; Liao, H. Y. M.; Fan, K. C.; Ko, M. T.; and Hung, Y. P., 1997. “Image Registration Using a New Edge-Based Approach. Computer Vision and Image Understanding”, Vol.67, No.2, pp. 112-130.

Hyung-Tae, K., Young-II, K., Woo-Sug., and Jeong-Hun S., 2000. “The Fusion of Photogrammetric Image and Laser Scanning Data”. In Proceeding of ASPRS Annual Conference on “The Imagine and GeoSpatial Information Society”, Washington, DC, 22-26 May. Published on CD-ROM.

Ingensand, H., Ryf, A., and Schulz, T., 2003. “Performances and Experiences in Terrestrial Laser Scanning”. In: Proceeding, Optical 3D-Measurement Techniques VI, Zürich, 2003, Switzerland, pp 236-244.

Ioannidis, C., Tsakiri, M., and Soile, S., 2004. “Rapid 3D Recording of Archaeological Sites Found During Construction Development”. Proceedings, International Workshop on Archaeological Surveys and Recording Methods, Greece , 22-27 May, 15p.

Karara, H. M., 1980. Non-Topographic Photogrammetry. In: *Manual of Photogrammetry American Society Photogrammetry*. New York, Falls Church, pp. 785-882.

Karras, G. E., and Mavrommati, D., 2001. “Simple Calibration Techniques for Non-metric Cameras”. CIPA International Symposium, Potsdam, September 18-21, pp. 1-8.

Kern. F., 2001. “Suplementing Laser Scanner Geometric Data with Photogrammetric Images for Modelling”. Proceedings, XVIII International Simposium on Survey Documentation of Historic Buildings-Monuments-Sites-Traditional and Decay Methods, Postdam, CIPA, 18-21 September, pp. 454-461.

King, B. A., 1993. “Methods for the Photogrammetric Adjustments of Bundles of Constrained Stereopairs”. PhD Thesis , Department of Civil Engineering and Surveying, The University of Newcastle, New South Wales, Australia, 220 p.

Knyaz, V.A., 2002. “Accurate Photorealistic Texture Mapping for Metric 3D Models”, Nizhny Novgorod, Russia, September 16 - 21, 2002.

Kuglin, C. D., and. Hines, D. C.,1975. “The Phase Correlation Image Alignment Method”, In Proceeding, IEEE 1975 International Conference on Cybernetics and Society, New York, pp. 163-165.

Kurazume, R., Nishino, K., Zhang, Z., and Ikeuchi, K., 2002. “Simultaneous 2D Images and 3D Geometric Model Registration for Texture Mapping Utilizing Reflectance Attribute”. The 5th Asian Conference on Computer Vision (ACCV2002), Melbourne, Australia, pp. 99-106.

Kurazume, R., Wheeler, M. D., and Ikeuchi, K., 2001. “Mapping Textures on 3D Geometric Model Using Reflectance Image”. Data Fusion Workshop in IEEE International Conference on Robotics and Automation. http://fortune.is.kyushu-u.ac.jp/~kurazume/papers/ICRA01-workshop.pdf ((Accessed on 20 June 2005).

Läbe, T., and Ellenbeck, K. H., 1996. “3D-Wireframe Models as Ground Control Points for the Automatic Exterior Orientation, In Proceedings, International Archives for Photogrammetry and Remote Sensing, Part B2, Vol. 31, pp218-223.

Lacey, A. J., Pinitkarn, N., and Thacker, N. A., 2000. “An Evaluation of the Performance of RANSAC Algorithms for Stereo Camera Calibration”. Electronic Proceedings of the Eleventh British Machine Vision Conference University of Bristol. 11-14 September 2000.
http://www.bnva.ac.uk/bmvc/2000/papers/p65.pdf.

Lee, Y., Habib, A., and Kim, K., 2002. “A Study on Aerial Triangulation from Multisensor Imagery”. In Proceedings, the International Symposium on Remote Sensing (ISRS), Sokcho, Korea, (October 30 – November 1), Vol.19, No.3, pp.255-261.

Lemmens, M. J. P. M., and van den Heuvel F. A., 2001. “Detailed and Accurate 3D Object Models from 3D Close-range Laser Mapping Syst”. In: GIM, Vol. 15, No.1, pp. 30-33.

Lhuillier, M., and Quan, L., 2000. “Robust Dense Matching Using Local and Global Geometric Constraints”. ICPR 2000, pp. 1968-1972.

Li, H., Manjunath, B. S., and Mitra, S. K.,1995. “Multisensor Image Fusion Using the Wavelet Trasnformation”. Graphical Models and Image Processing, vol. 57, pp. 235-245.

Liang-chien, C., Tee-ann, T., Yi-chen, S., Yen-chung, L., and Jiann-yeou, R., 2004. “Fusion of Lidar Data and Optical Imagery for Building Modeling”**.** Proceedings, XXth ISPRS congress, Istanbul, Turkey, 12-23 July 2004, paper 732.

Lichti, D., 2004. "A Resolution Measure for Terrestrial Laser Scanners". Proceedings, XXth ISPRS Congress, Istanbul, Turkey, 12-23 July, paper 552.

Lichti, D. D., and Gordon, S. J., 2004. “Error Propagation in Directly Georeferenced Terrestrial Laser Scanner Point Clouds for Cultural Heritage Recording”. FIG Working Week 2004, Athens, Greece, May 22-27. http://www.fig.net/pub/athens/papers/wsa2/WSA2_6_Lichti_Gordon.pdf.

Lichti, D. D., and Harvey, B. R., 2002a. “An Investigation into the Effects of Reflecting Surface Material Properties on Terrestrial Laser Scanner Measurements”. Geomatics Research Austrlasia, No. 76, pp. 1-22.

Lichti, D. D., Gordon, S. J., and Stewart, M. P., 2002b. “Ground-based Laser Scanners: Operation, Systems and Application”. Goematica. Vol. 56, No.1, pp. 21-33.

Lichti, D. D., Stewart, M. P., Tsakiri, M., and Snow, A. J., 2000. "Benchmark Tests on Three-Dimensional Laser Scanning System". Geomatics Research Australasia, No. 72, pp. 1-24.

Malian, A., Heuvel van den, F. A., and Azizi, A., 2002. "A Robust Photogrammetric System for Wound Measurement." International Archives of Photogrammetry and Remote Sensing, Vol. 34, Part 5, pp. 264 – 269.

Mason, S., 1995. "Conceptual Model of the Convergent Multistation Network Configuration Task". Photogrammetric Record, 15(86): 277-299.

McGlone, Chris, 1996. "Bundle Adjustment with Geometric Constraints for Hypothesis Evaluation". In: International Archives of Photogrammetry and Remote Sensing, XXXI, B3, Vienna, Austria July. pp 529-534.

Mikhail, E. M., Bethel, J. S., and McGlone, J. C., 2001. *Introduction to Modern Photogrammetry*. John Wiley & Sons.

Morgan, M., and Habib, A., 2002. "Interpolation of Lidar Data and Automatic Building Extraction". ACSM-ASPRS 2002 Annual Conference Proceedings. http://www.geomatics.ucalgary.ca/~habib/r_published_work.html.

Naai-Jung, S., Ming-Chang, W., and Kunz, J., 2004. "The Inspections of As-built Construction Records by 3D Point Clouds". Center for Integrated Facility Engineering (CIFE),Working Paper, August, Stanford University, paper 090.

Nakagawa M., and Shibasaki R., 2001. "Study on City Model with Image and Laser Data". Proceeding of Asian Conference on Remote Sensing, Singapore, 5-9 November. http://www.crisp.nus.edu.sg/~acrs2001/pdf/255masa.pdf.

Nakagawa, M., Shibasaki, R., and Kagawa, Y., 2002, “Fusion Stereo Linear CCD Image and Laser Range Data for Building 3D Urban Model”. IAPRS, Vol.34, Part 4, pp. 200-211.

Nickels, K. M., Castano, A., Cianci., C., 2003. “Fusion of Lidar and Stereo Range for Mobile Robots”. Proceedings, the 11th International Conference on Advanced Robotics, Coimbra, Portugal, June 30 - July 3, 6p.

Nobel, J., 1998. “Finding Corners”. Image and Vision Computing, Vol.6 pp 122-128.

Ono N., Tonoko N., and Sato K., 2000. “A Case Study on the Landslide by the 3D Laser Scanner”. International Archives of Photogrammetry and Remote Sensing. Vol. XXXIII, Part B5, pp. 593-598.

Pfeiffer, N., and Rottensteiner, G., 2001. “The Riegl Laser Scanner for the Survey of the Interiors of Schönbrunn Palace. Fifth Conference on Optical 3-D Measurement Techniques. Vienna, 4-6 October. http://www.ipf.tuwien.ac.at/ publications/np_vienna_terrestrialls_2001.pdf.

PhotoModeler., 2005. Photogrammetry Corner. http://www.photomodeler.com /cservicepc.html (Accessed on 15 June 2005).

Pollefeys M., Koch R., Verguawen M., and Van Gool L., 2001. “Automated Reconstruction of 3D Scenes from Sequences of Images”. ISPRS Journal of Photogrammetry and Remote Sensing. Vol. 55, No.4, pp. 251-267.

Phol, C., and van Genderen, J. L., 1998. “Multisensor Image Fusion in Remote Sensing: Concepts, Methods and Applications”. International Journal of Remote Sensing, 19(5), pp. 823-854.

Pulli, K., and Shapiro, L., 2000. “Surface Reconstruction and Display from Range and Colour Data”. *Graphical Models and Image Processing,* 62, pp. 165-201.

Qi, Z. R., 2004. **“**A New Data Fusion Method for Improving Cbers-1 Irmss Images Based on CCD Image Multi-source Imagery”. Proceedings, XXth ISPRS Congress, Istanbul, Turkey, 12-23 July, paper 910.

Rahman, Z., Jobson, D. J., and Woodell, G. A., 2002. “Multisensor Fusion and Enhancement Using the Retinex Image Enhancement Algorithm”. SPIE International Symposium on AeroSense, Proceedings of the Conference on Visual Information Processing XI. http://dragon.larc.nasa.gov/retinex/background/pubabs /papers/orl2002.pdf.

Rajan D., Chaudhuri, S., 2001. Generalized Interpolation and its Application in Super-Resolution Imaging. Image and Vision Computing 19, pp. 663–679.

Reddy, S. B., and Chatterji, B. N., 1996. “An FFT-based Technique for Translation, Rotation, and Scale-Invariant Image Registration”. IEEE Transaction on Image Processing, Vol. 5, No.8, pp. 1266-1271.

Reiss P., 2002. “Combination of Lidar, Digital Photogrammetry and Terresrtrial Survey to Generate High-Quality DEMs”. IAPRS and SIS. Vol 34, Part 4, pp. 526-531.

Remondino, F., and Borlin, N., 2004. “Photogrammetric Calibration of Sequences Acquired with a Rotating Camera”. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, part 5/W16. ISPRS Panoramic Photogrammetry Workshop, Dresden, Germany, 19-22 February. http://www.photogrammetry.ethz.ch/general/persons/fabio /isprs_dresden.pdf.

Rockett, P. I., (2003). “Performance Assessment of Feature Detection Algorithms: a Methodology and Case Study on Corner Detectors”. IEEE Transactions on Image Processing, Vol. 12, No. 12, pp 1668-1676.

Rothfeder, J. L., Feng, S., and Rath, T. M., 2003. “Using Corner Feature Correspondences to Rank Word Images by Similarity”. Proceeding, the Workshop on Document Image Analysis and Retrieval (DIAR). http://ciir.cs.umass.edu/ pubfiles/mm-44.pdf.

Rottensteiner, F., and Jansa, J., 2002, “Automatic Extraction of Building from LIDAR Data and Aerial Images”, IAPRS, Vol.34, Part 4, pp. 295-301.

Santala, J., and Joala, V., 2003. “On the Calibration of a Ground-based Laser Scanner”. FIG Working Week 2003, Paris, France, April 13-17. http://www.fig.net/pub/fig_2003/TS_12/TS12_4_Santala_Joala.pdf.

Saroglu, E., Bektas, F., Musaoglu, N., and Goksel C., 2004. “Fusion of Multisensor Remote Sensing Data: Assessing the Quality of Resulting Images”. Proceedings, XXth ISPRS Congrss, Istanbul, Turkey, 12-23 July, paper 575.

Sasagawa, T., Babu, M. B., Ozawa, A., and Tachibana, K., 2004. “Capability for Data Fusion by Airborne Sensing”. Proceedings, XXth ISPRS Congress, Istanbul, Turkey, 12-23 July, paper 901.

Schenk, T., and Csatho, B., 2002. “Fusion of LIDAR and Aerial Imagery for a More Complete Surface Description”. In: International. Archives of Photogrammetry and Remote Sensing 34(3A), pp. 310-317.

Schenk T., 1999. *Digital Photogrammetry*. Terrascience, Laurelville, Ohio, pp 355-385.

Schiewe, J., 2004**. "**Fusion and Perceptual Organisation of Features from Multisensor Data: General Concepts and New Developments"**.** Proceedings, XXth ISPRS Congress, Istanbul, Turkey, 12-23 July, paper 605.

Schiewe J., 2000. "Improving the Integration of Digital Surface Models". International Archives of Photogrammetry and Remote Sensing. Vol. 32, Part B3, pp. 807-813.

Schiewe J., 1999. "An advanced Technique for Pixel-based Multisensor Data Integration"**.** In, Proceedings of the ISPRS-workshop Sensors and Mapping from Space 1999, Hannover, Germany. Published on CD-ROM.

Schmid, C., Mohr, R., and Bauckhage, C., 2000. "Evaluation of Interest Point Detectors". In: International Journal of Computer Vision, 37(2), pp 151-172, 2000.

Singh R., Chapman D. P., and Atkinson K. B., 1997. "Digital Photogrammetry for Automatic Close-range Measurement of Textureless and Featureless Objects". The Photogrammetric Record. Vol. XV, No.89, pp 691-702.

Smith, P., Sinclair D., Cipolla R., and Wood K., 1998. "Effective Corner Matching". On-Line Proceedings of the Ninth British Machine Vision Conference. http://www.bmva.ac.uk/bmvc/1998/papers/d125/h125.htm.

Smith, S. L., Holland, D. A., and Longley, P. A., 2003a. "Interpreting Interpolation: The Pattern of Interpolation Errors in Digital Surface Models Derived from Laser Scanning Data". Proceeding, GIS Research UK 11th Annual Conference, London. http//www.casa.ucl.ac.uk/working_papers/paper66.pdf.

Smith, S. L., Holland, D. A., and Longley, P. A., 2003b. "The Effect of Changing Grid Size in the Creation of Laser Scanner Digital Surface Models". Proceedings.

Geocomputation. On-Line Proceedings of the 7th International Conference on GeoComputation. http://www.geocomputation.org/2003/Papers/Smith_Paper.pdf.

Staiger, R., 2003. Terrestrial Laser Scanning – Technology, Systems and Applications. 2nd FIG Regional Conference, Marrakech, Morocco, December 2-5.

Staiger R., 2002. "Laser Scanning in An Industrial Environment". FIG XXII International Congress, Washington, D.C, USA, pp 1-8.

Stamos, I., 2001. Geometry and Texture Recovery of Scenes of Large Scale: Integration of Range and Intensity Sensing. PhD Thesis, Graduate School of Arts and Science, Columbia University.

Stamos I., Allen P., 2000. "3D Model Construction Using Range and Image Data". Computer Vision and Pattern Recognition, pp. 537-544.

Styladis, A. D., Karagiannidis C. J., Zestas, N. C., and Chatzara, K. N., 2003. "Pose Determination from a Single Image in a Controlled CAD Environment", Journal of WSCG, Vol. 11 (1). http://wscg.zcu.cz/wscg2003/Papers_2003/E11.pdf.

Sowmya A., and Trinder J., 2000. "Modelling and Representation Issues in Automated Feature Extraction from Aerial and Remote Sensing". ISPRS Journal of Photogrammetry and Remote Sensing. Vol. 55, No.1, pp 34-47.

Thiyagarajan, R., 2003. "Use of 3D Laser Scanning and Close-range Photogrammetry for Ship Building". Proceedings, Ship Production Track of the World Maritime Conference, San Francisco, 17-20 October. http://www.skibsteknisk selskab.dk/download/WMTC/F8(P81).pdf.

Tissainayagam, P., and Suter, D., 2004. “Assessing the Performance of Corner Detectors for Point Feature Tracking Applications”. Image and Vision Computing 22, pp. 663–679.

Vajdic, S. M., Brooks, M. J., Downing, A., and Katz, H. E., 1995. “AI and Medical Imagery: Strategy and Evaluation of Inexact Relational Matching”. Complexity International, Vol. 2. http://journal-ci.csse.monash.edu.au/ci/vol02 /vajdic/vajdic.html.

Vosselman, G., 2002. “Fusion of Laser Scanning Data, Maps and Aerial Photographs for Building Reconstruction”. International Geoscience and Remote Sensing Symposium, 2002, 24-28 June, Toronto, Canada, Published on CD-ROM.

Walstra. J., Chandler, J. H., Dixon, N., and Dijkstra, T. A., 2004. Time for Change – Quantifying Landslide Evolution using Historical Aerial Photographs and Modern Photogrammetric Methods. Proceedings, XXth ISPRS Congress Istanbul, Turkey, 12-23 July, paper 395.

Wang, X., and Clarke, T. A., 1998. “Separate Adjustment of Close-range Photogrammetric Measurements”**.** ISPRS Vol. XXXII**,** Part 5. pp. 177-184.

Weinhaus, F. M., Devarajan, V., 1997. “Texture Mapping 3D Models of Real-World Scenes”. ACM Computing Surveys, Vol. 29, No. 4. pp. 325-365.

Xie, H., Hicks, N., Keller, R. G., Huang, H., and Kreinovick, V., 2000. Automatic Image Registration Based on FFT Algorithm and IDL/ENVI, ICORG-2000 International Conference on Remote Sensing And GIS/GPS Proceedings. Dec. 1-4, Hyderabad, India. Vol. 1, pp. 397-402.

Yong F., Benyi, C., Wensong, H.,and Hong, C., 2000. “DEM Generation from Multisensor SAR Images”. International Archieves of Photogrammetry and Remote Sensing. Vol. XXXIII, Part B1, pp. 110-116.

Yu, Y., and Ferencz, A., 2001. “Extracting Objects from Range and Radiance Images”. IEEE Transactions on Visualization and Computer Graphics, Vol. 7, No. 4, pp. 351-364.

Zitova, B., and Flusser, J., 2003. “Image Registration Methods: A Survey”. Image and Vision Computing, 21 (2003), 11, pp 977-1000.

APPENDIX A1

REAL CAMERA 2D IMAGES

FigureA1.1: Real Camera 2D Images of Data Set A.

Figure A1.2: Real Camera 2D Images of Data Set B.

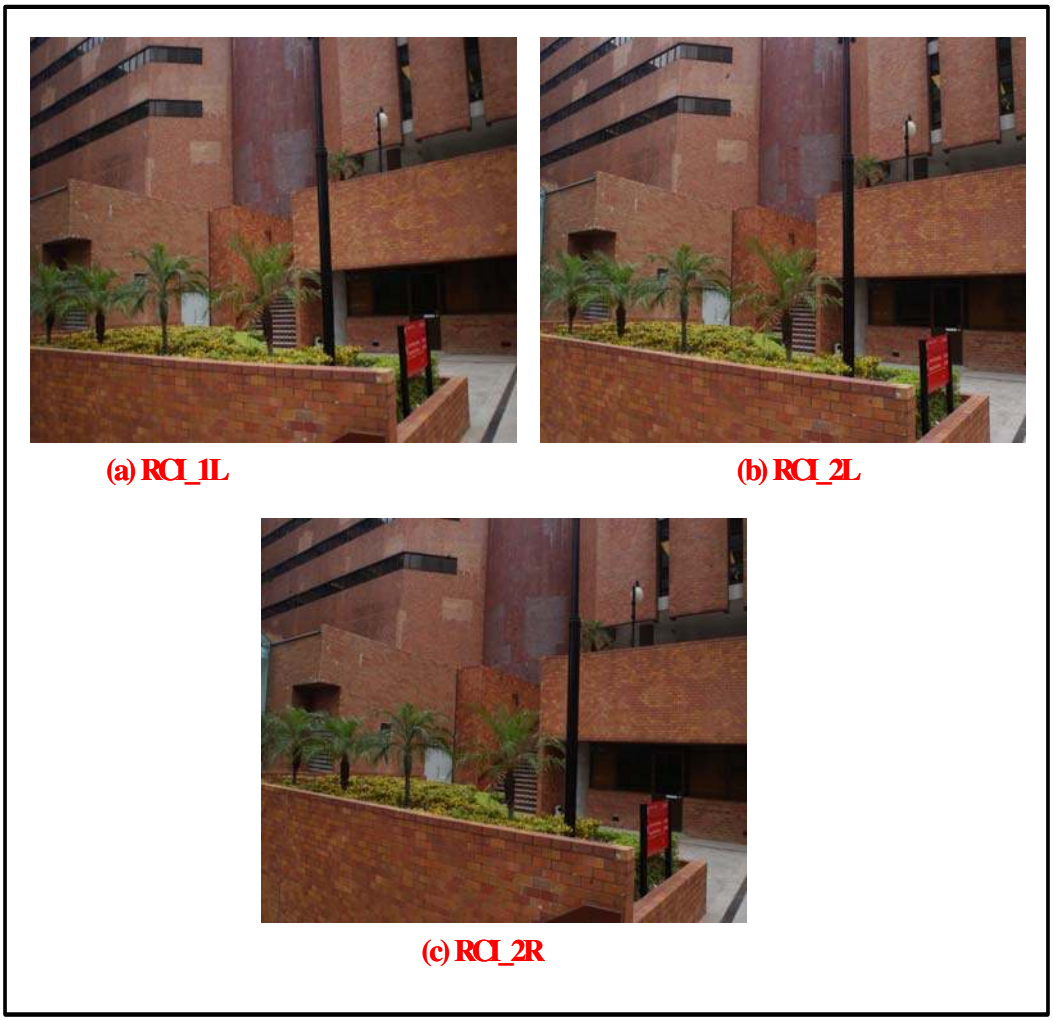


Figure A1.3: Real Camera 2D Images of Data Set C.

APPENDIX A2

FFT-BASED MULTISENSOR IMAGE REGISTRATION

A2: Introduction

This section presents the necessary theory for the pairwise image registration using Fourier translation, rotation and scaling properties

A2.2: The Theory

It is known that if two images f_1 and f_1 differ only by shift (x_0, y_0), i.e.,

$$f_1(x, y) = f_2(x - x_0, y - y_0), \tag{A2.1}$$

then by the shifting property of Fourier Transforms:

$$F_2(\omega_1, \omega_2) = F_1(\omega_1, \omega_2)e^{-j(\omega_1 x_0 + \omega_2 y_0)} \tag{A2.2}$$

The translational offsets (x_o, y_o) can be recovered by locating the impulses associated with the inverse transform of the cross-power spectrum of the two images (Eustice *et al*, 2002). The ratio R (in Equation A2.3), is the cross-power

spectrum of two images f_1 and f_2 with Fourier transforms F_1 and F2 is defined as (Eustice *et al*, 2002; Reddy & Chatterji, 1996; Xie *at* al, 2000):

$$R = \frac{F_1(\omega_1, \omega_2) * \text{conj}(F_2(\omega_1, \omega_2))}{\text{abs}(F_1(\omega_1, \omega_2)) * \text{abs}(F_2(\omega_1, \omega_2))} \quad \text{(A2.3)}$$

Where conj is the complex conjugate, abs is absolute value. By taking an inverse Fourier transform of R, the position with the maximum absolute value can be found when everywhere else is approximately zero. This position (x_o, y_o) is the displacement that is needed to optimally register the images. The same property can be exploited for images, which are rotated and scaled by representing them in a coordinate system where scale and rotation appear as shifts (Eustice *et al*, 2002). For example, when $f_2(x, y)$ is a rotated version of $f_1(x, y)$:

$$f_2(x, y) = f_1(x\cos\theta_0 + y\sin\theta_0, -x\sin\theta_0 + y\cos\theta_0) \quad \text{(A2.4)}$$

Their Fourier Transforms are related by:

$$F_2(\omega_1, \omega_2) = F_1(\omega_1\cos\theta_0 + \omega_2\sin\theta_0, -\omega_1\sin\theta_0 + \omega_2\cos\theta_0) \quad \text{(A2.5)}$$

Converting the magnitude of the Fourier Transform from rectangular coordinates to log polar coordinates makes is possible to represent rotation as shifts (Eustice *et al*, 2002):

$$M_2(\rho,\theta) = M_1(\rho,\theta-\theta_0) \tag{A2:6}$$

$$\rho = \sqrt{\omega_1^2+\omega_2^2}, \quad \theta = \arctan(\omega_1/\omega_2) \tag{A2:7}$$

Similarly, when two images are related by a scale factor, *a,* then their Fourier Transform is related by:

$$F_2(\omega_1,\omega_2) = \frac{1}{a^2}F_1(\omega_1/a,\omega_2/a) \tag{A2:8}$$

Taking the logarithm of the frequency results in the scale appearing as a shift (Eustice *et al*, 2002):

$$F_2(\log\omega_1,\log\omega_2) = \frac{1}{a^2}F_1(\log\omega_1-\log a,\log\omega_2-\log a) \tag{A2.9}$$

Combining all these properties, it is noticed that the magnitudes of the two translated, scaled, rotated images are related:

$$M_2(\rho,\theta) = M_2(\rho,\theta-\theta_0) \tag{A2:10}$$

After taking the log of the radius, rotation and scale are now both represented as shifts:

$$M_2(\rho,\theta) = M_2(\log\rho-\log a,\theta-\theta_0) \tag{A2:11}$$

A2.3: Results of the FFT-based Registration

Table A2.1: Parameters for Registration using the FFT for Data Set A

	Scale	Rotation (θ°)	Translation (x_o, y_o)	P_Value
Model_1	0.997	179.673	-41,13	0.117
Model_2	0.997	180.327	21, 105	0.136
Model_3	0.890	177.709	23, 86	0.087
Model_4	0.904	2.945	68, 25	0.066

Table A2.2: Parameters for Registration using the FFT for Data Set B

	Scale	Rotation (θ°)	Translation (x_o, y_o)	P_Value
Model_1	1.016	0.194	229, -3	0.222
Model_2	1.002	0.194	19, -9	0.157
Model_3	1.031	180.972	-17, 23	0.141
Model_4	1.0119	179.806	229, 148	0.156

Table A2.3: Parameters for Registration using the FFT for Data Set C

	Scale	**Rotation** (θ°)	**Translation** ($x_{o,}\ y_{o}$)	**P_Value**
Model_1	0.997	270.000	53 ,33	0.033
Model_2	0.997	270.000	-29, 9	0.032
Model_3	0.997	2.571	-29, 6	0.031
Model_4	0.997	0.000	-34, -5	0.028

APPENDIX A3

RESULTS WITH THE HARRIS CORNER DETECTOR

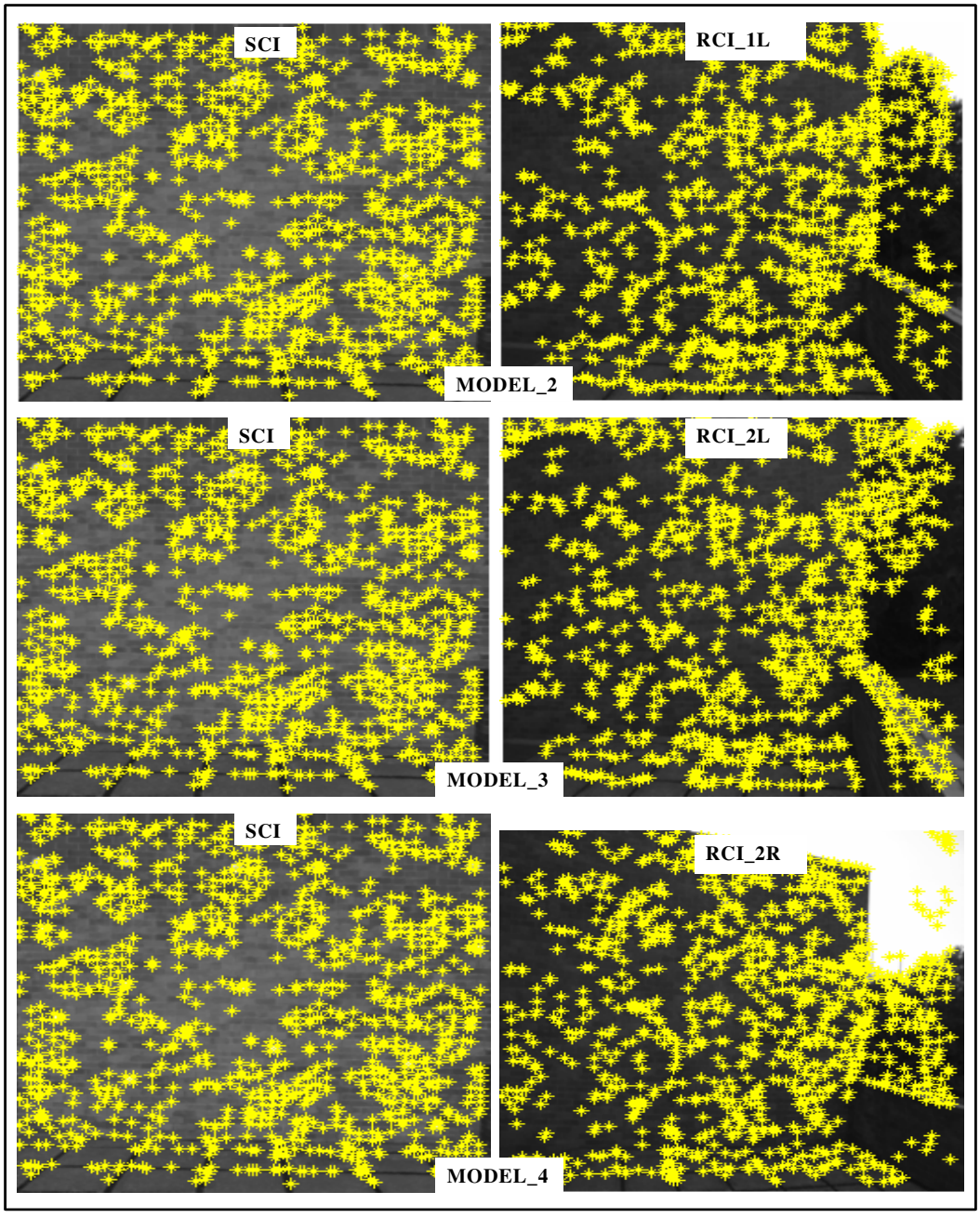


Figure A3.1: Example of Corner Detected with Harris Corner Detector on Data Set A.

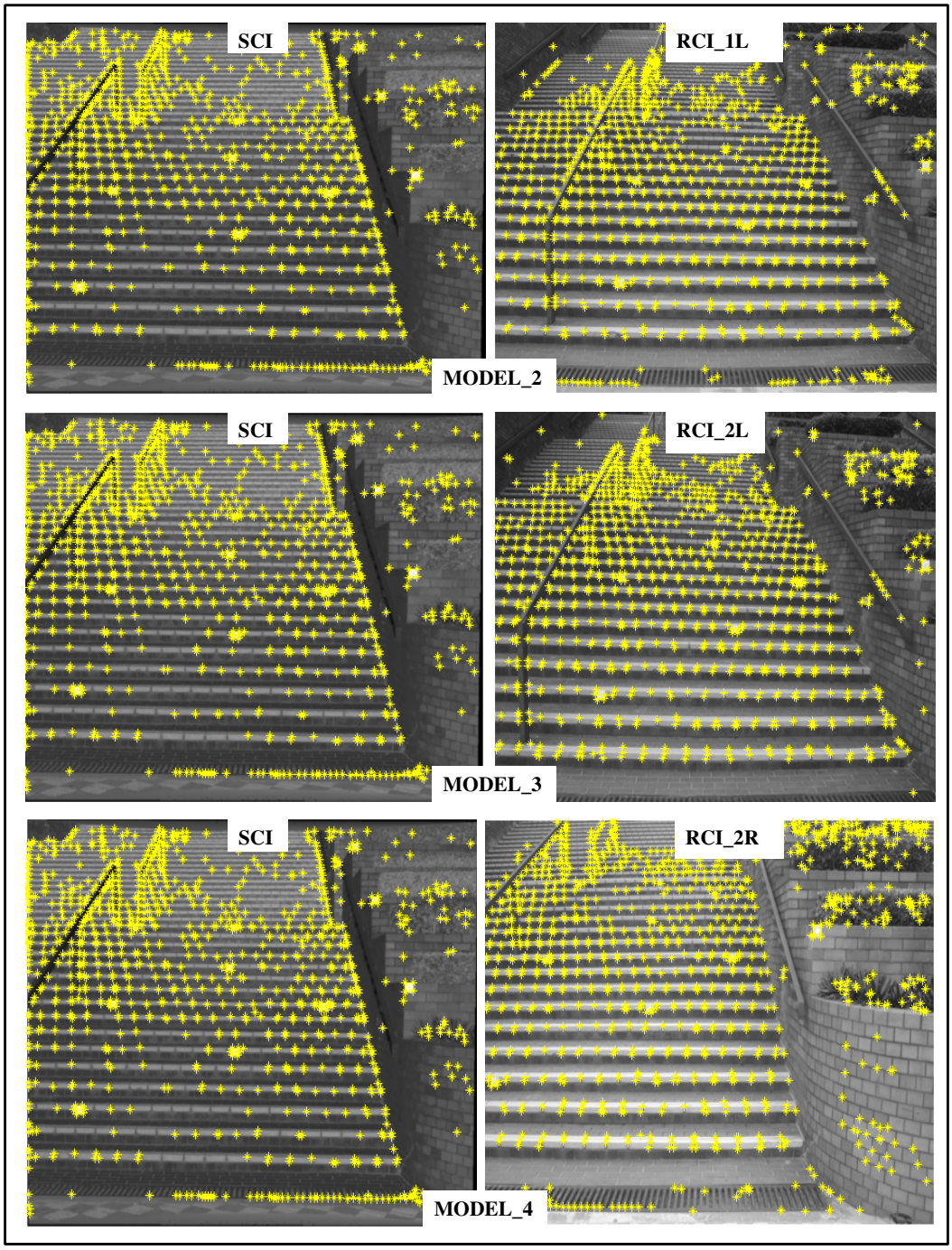


Figure A3.2: Example of Corner Detected with Harris Corner Detector on Data Set B.

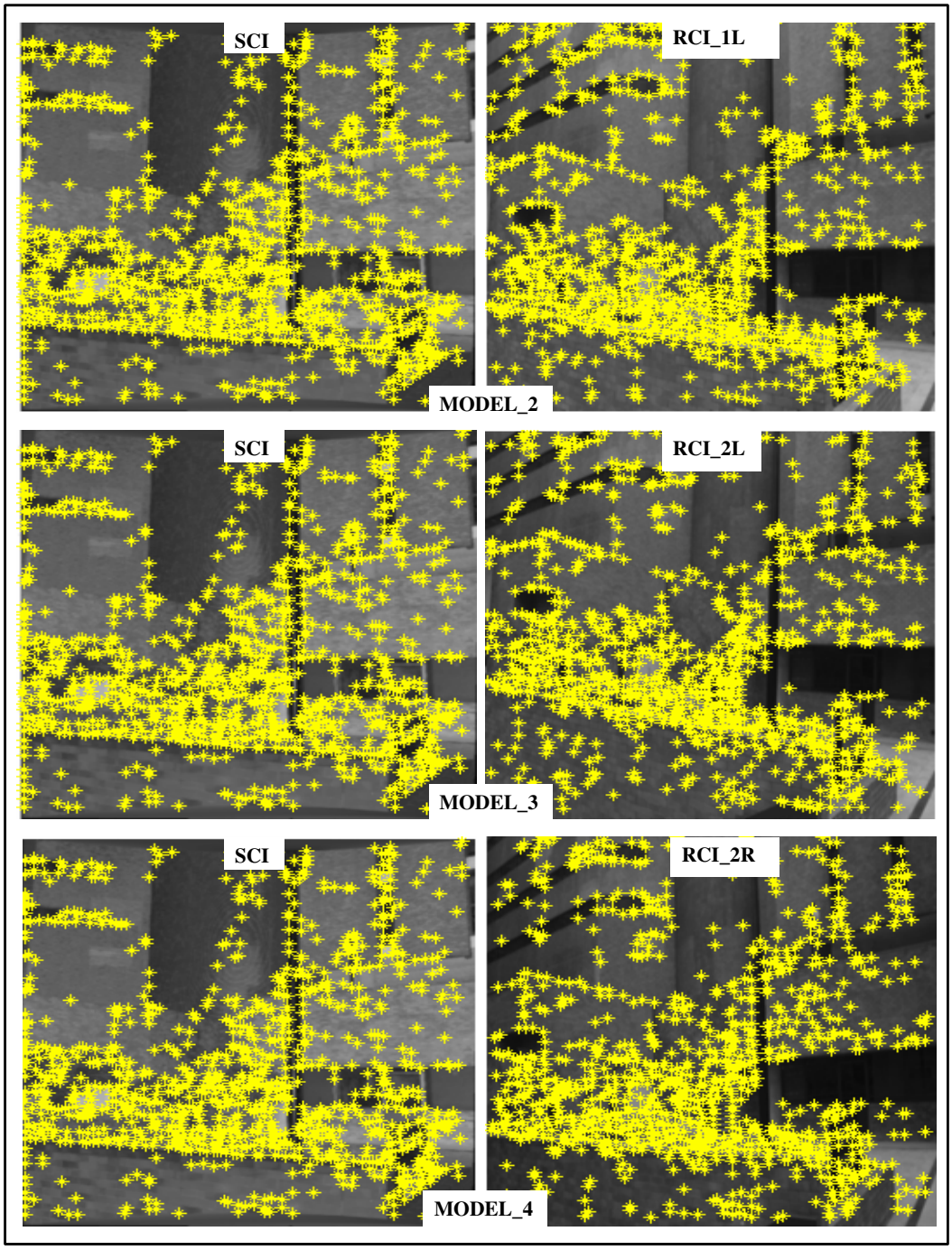


Figure A3.3: Example of Corner Detected with Harris Corner Detector on Data Set C.

APPENDIX A4

RESULTS OF THE AUTOMATIC HYBRID MATCHING

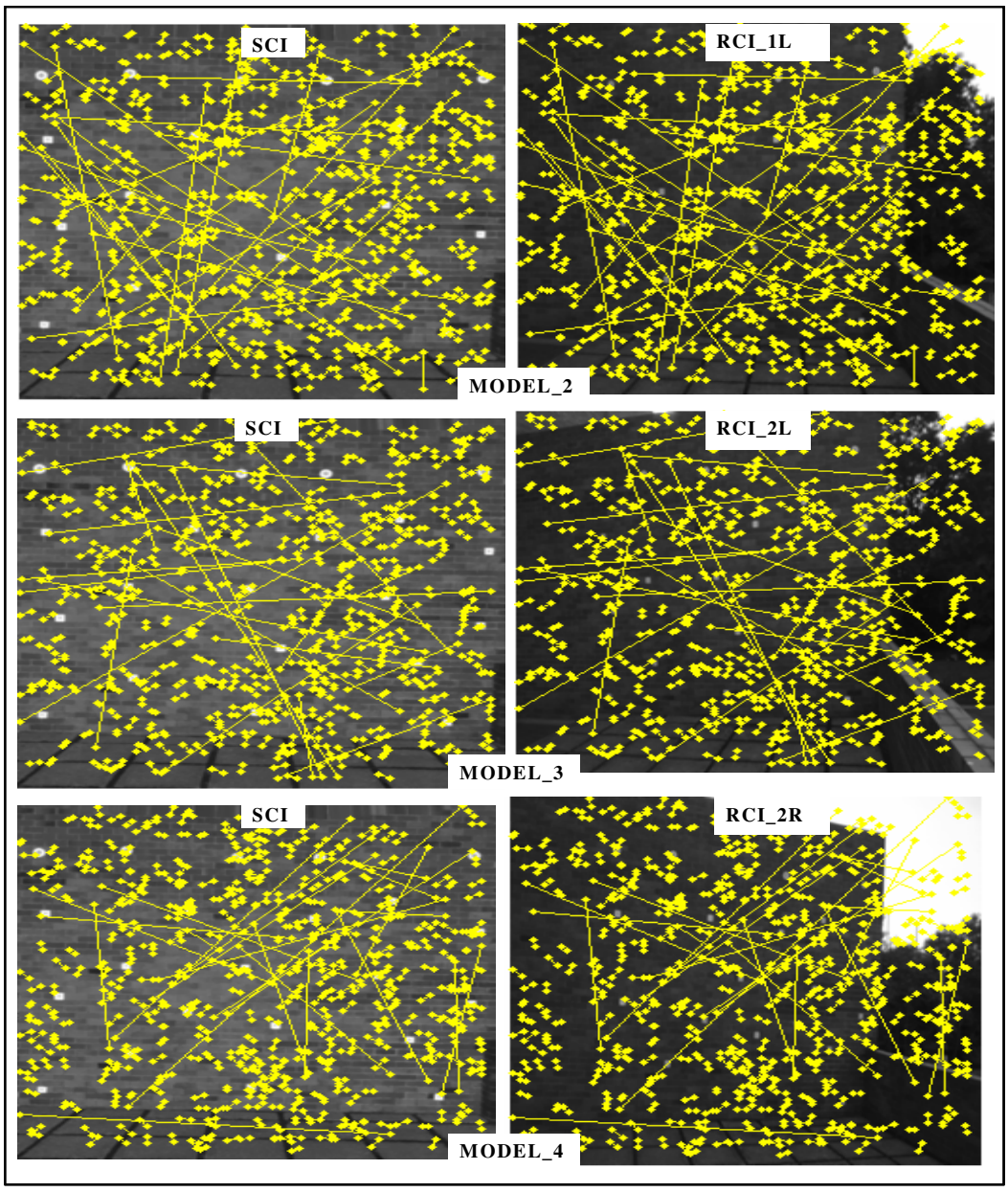


Figure A4.1: Detected Correspondences of Model_1 of Data Set A with ZNCC.

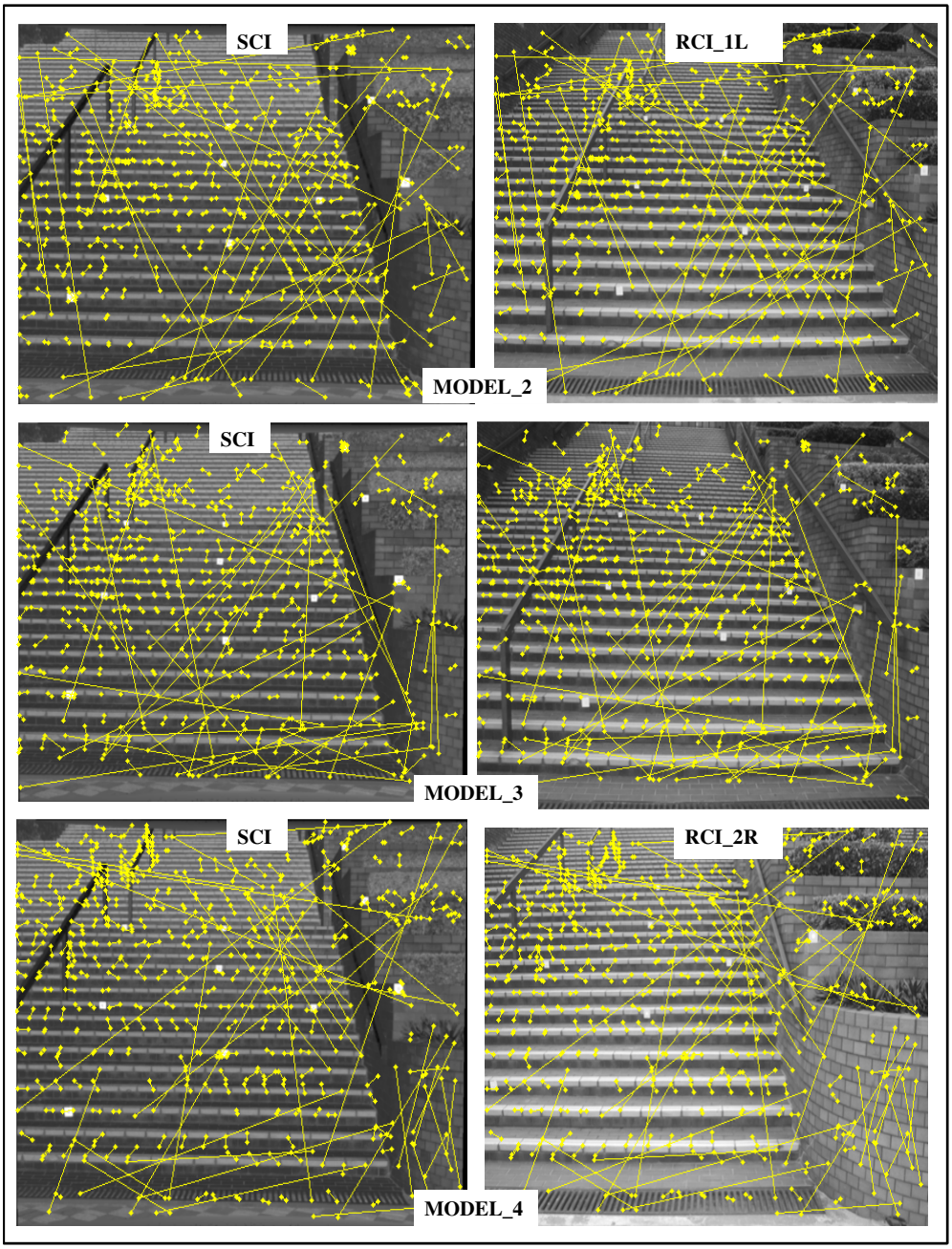


Figure A4.2: Detected Correspondences of Model_1 of Data Set B with ZNCC.

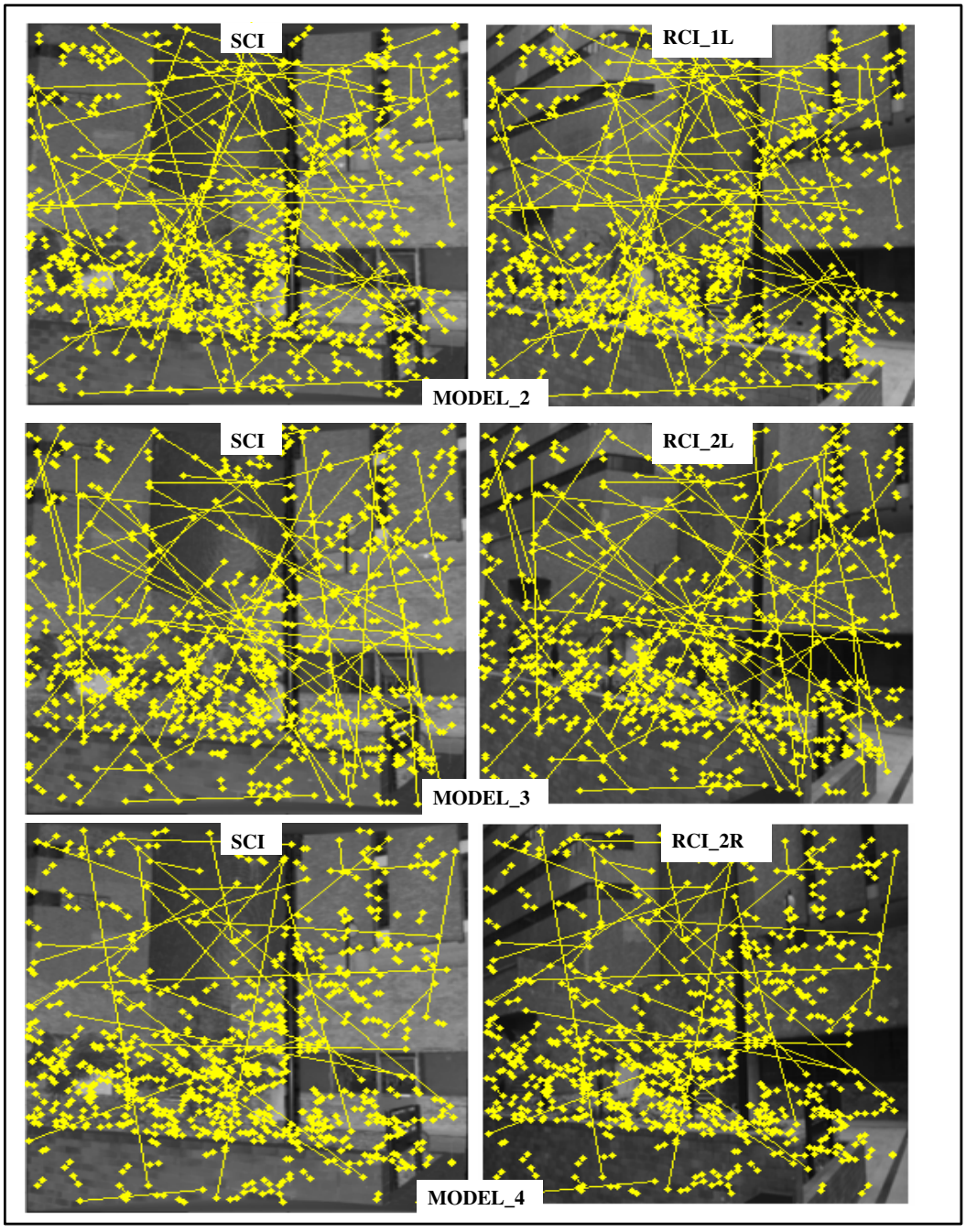


Figure A4.3: Detected Correspondences of Model_1 of Data Set C with ZNCC.

APPENDIX A5

RESULTS OF THE FINAL VERIFIED DETECTED CORRESPONDENCES (RANSAC INLIERS)

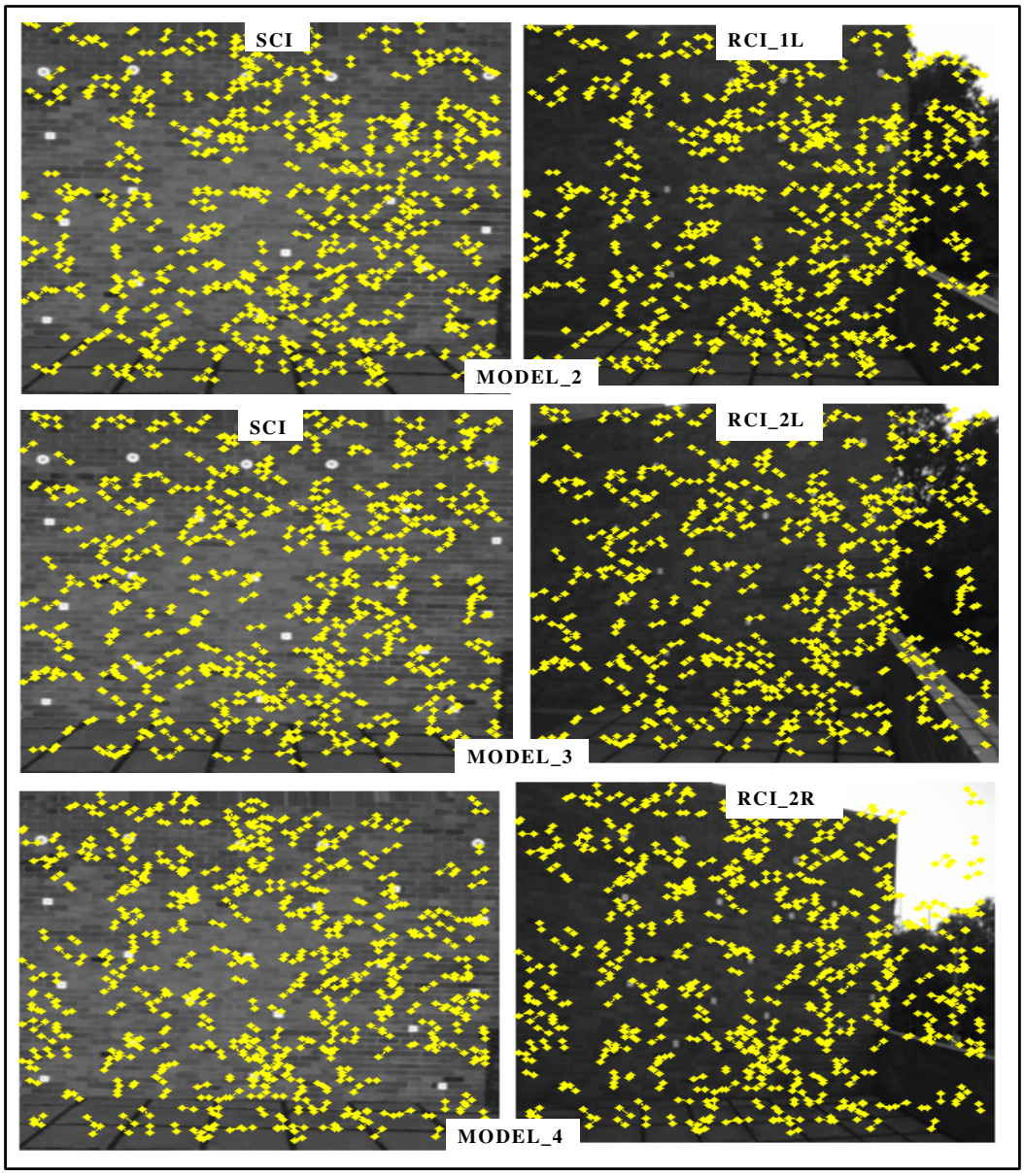


Figure A5.1: The Final Verified Detected Correspondences (RANSAC inliers) of Model_1 of Data Set A.

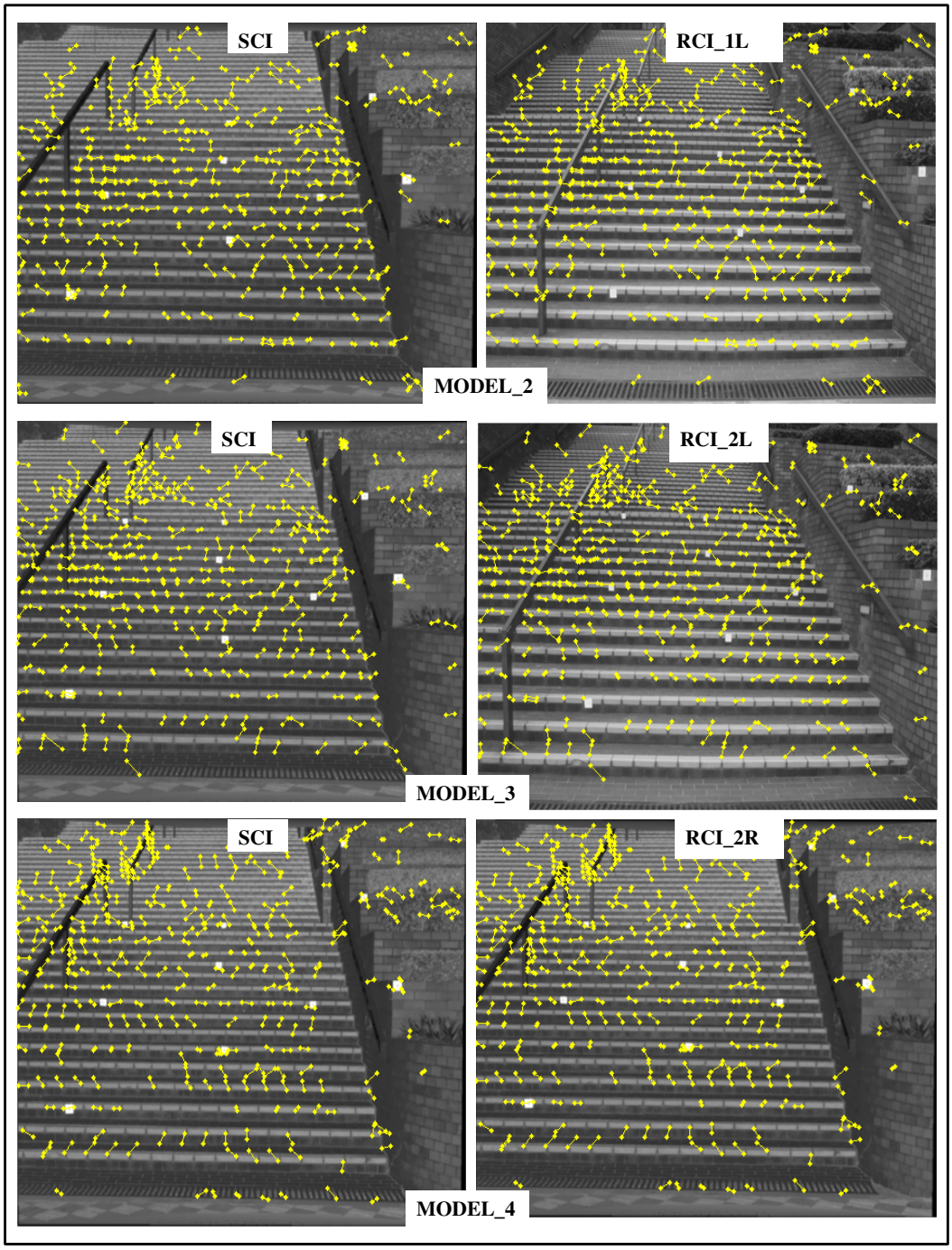


Figure A5.2: The Final Verified Detected Correspondences (RANSAC inliers) of Model_1 of Data Set B.

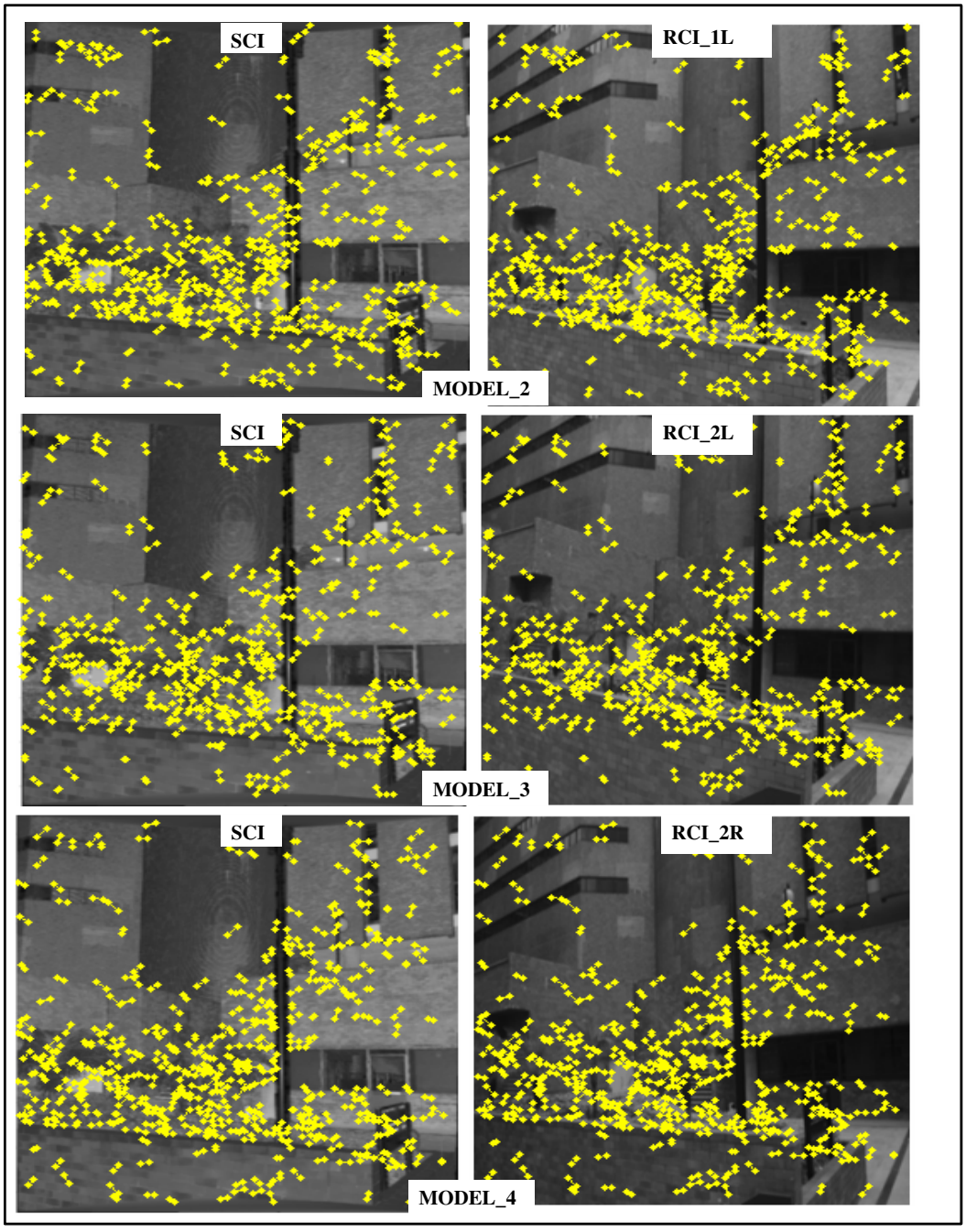


Figure A5.3: The Final Verified Detected Correspondences (RANSAC inliers) of Model_1 of Data Set C.

APPENDIX A6

OUTLINE OF SOFTWARE ENVIRONMENTS, TOOLS DEVELOPED AND READY-MADE TOOLS USED

A6.1: Introduction

This section presents the list of software used, the tools developed and ready-made tools used in the thesis.

A6.2: Outline of Software Environments

1. Cyra's Cyclone software was used to capture the 3D point clouds, preprocessing the point clouds (data filtering) and to export the point clouds as ASCII format. Also, this software was used to transform the 3D point clouds into an intermediate surface-based representation.

2. 3D Mapper StereoMaker software was used for the implementation of the manual exterior orientation measurements.

3. Autodesk 3Ds Max software was used to merge the surface-based representation and the texture image into a common representation (photorealistic models)

A6.3: Tools Developed

The Matlab programming language was used to implement the following:

- Read and process the 3D point clouds
- Create the synthetic camera images
- Image pyramids
- Extraction of points of interest using the Harris corner detector.
- Correspondence matching using zero mean normalized cross correlation algorithm
- Refines the initial corresponding points by using a robust estimator such as RANSAC

A6.4: Ready-made Tools Used

The bundle adjustment as presented in Simultaneous Photogrammetric and Geodetic Adjustment (SPGA) developed by Fraser (1983) was used.

APPENDIX A7

DATA SETS USED AND MATLAB CODES DEVELOPED FOR THE MULTISENSOR FUSION PROCESS

A7.1 Introduction

This section describes the data sets used, Matlab codes developed for the multisensor fusion and the results obtained. These listings can be found on the DVD disk at the end of this thesis.

A7.2 Data Sets Used and Results of Pre-Processing

A7.2.1 Raw Cyrax Data (with imp File Extension)

The files are contained in \imp_file

Data Set_A.imp;

Data Set_B.imp

Data Set_C.imp

A7.2.2 ASCII files of the 3D Point Clouds (with pts File Extension)

The files are contained in \ASCII_data

Data Set_A.pts;

Data Set_B.pts

Data Set_C.pts

A7.2.3 Back Projected Points (after Collinearity Equations)

The files are contained in \Collin_data

Data Set_A.dat;

Data Set_B.dat

Data Set_C.dat

A7.2.4 Generated Synthetic Camera Images – SCIs

The files are contained in \SCI_image

SCI_A

SCI_B

SCI_C

A7.2.5 Real Camera Images – RCIs

The files are contained in \RCI_image:

Data Set A

RCI_1R,

RCI_1L

RCI_2L

RCI_2R

Data Set B

RCI_1R

RCI_1L

RCI_2L

RCI_2R)

Data Set C

RCI_1R,

RCI_1L

RCI_2L

RCI_2R

A7.3 Results of Automatic Measurements

A7.3.1Image Pyramid

The files are contained in \Pyramid_image:

Data_Set_A

Data_Set_B

Data_Set_C

A7.3.2 Feature Extracted with Harris Corner Detector

The files are contained in \Harris_Corner

Data Set A

- H11
- H22
- H33
- H44

Data Set B

- H11
- H22
- H33
- H44

Data Set C

- H11
- H22
- H33
- H44

A7.3.3 Correspondence Matching with ZNCC

The files are contained in \ZNCC

Data Set A

- ZNCC11

ZNCC22

ZNCC33

ZNCC44

Data Set B

ZNCC11

ZNCC22

ZNCC33

ZNCC44

Data Set C

ZNCC11

ZNCC22

ZNCC33

ZNCC44

A7.3.4 RANSAC

The files are contained in \RANSAC

Data Set A

R11

R22

R33

R44

Data Set B

R11

R22

R33

R44

Data Set C

R11

R22

R33

R44

A7.4 Results of Orientation Measurements

A7.4.1 Manual Exterior Orientation Results

The files are contained in \MO

Data Set A

Data Set B

A7.4.2 Automatic Exterior Orientation Results

The files are contained in \AO

Data Set A

Data Set B

Data Set C

A7.5 Results of Texture Mapping

The file is contained in \TM

A7.6 Matlab Codes for the Multisensor Fusion Process

A7.6.1 Code to Read Point Clouds, to Perform Backprojection and to Generate Synthetic Images

The file is contained in \matlab_code\read_dat.m

A7.6.2 Code to Extract Points of Interest, Correspondence Matching (with ZNCC), and RANSAC

The file is contained in \matlab_code\ransac.m